القياس النفسي والتربوي

التعريف بالقياس ومفاهيمه وأدواته
بناء المقاييس ومميزاتها
القياس التربوي

. تأليف

الدكتور محمد عبد السلام أحمد

ملتزمة الطبع والنشر
مكتبة النهضة المصرية
لأصحابها حسن محمد وأولاده
٩ شارع عدلي باشا بالقاهرة

القياس النفسي والتربوي

التعريف بالقياس ومفاهيمه وأدواته
بناء المقاييس ومميزاتها
القياس التربوي

تأليف

الدكتور محمد عبد السلام أحمد

ملتزمة الطبع والنشر
مكتبة النهضة المصرية
لأصحابها حسن محمد وأولاده
٩ شارع عدلي باشا بالقاهرة

الطبعة الأولى

١٩٦٠

القـــاهرة

مطبعة لجنة التأليف والترجمة والنشر

مقدمة

موضوع هذا الكتاب هو أهم موضوعات علم النفس بل هو العماد الذى يقوم عليه علم النفس كعلم Science فلم يتحول علم النفس من ميدان الفلسفة إلى ميدان العلوم إلا عند ما أخذ الباحثون فيه بالمنهج العلمى ، ذلك المنهج الذى يعتمد أساساً على الملاحظة الموضوعية . فكما قال كارل بيرسون Karl Pearson العالم الرياضى المشهور « إن الأساس الذى تقوم عليه العلوم ليس طبيعة مادتها أو موضوعها ولكنه الأسلوب العلمى أو الطريقة العلمية التى تتبعها » . ويقاس تقدم أى علم من العلوم بدرجة الدقة التى يصل إليها فى القياس ، على أن عملية القياس فى علم النفس والتربية أعقد وأشق كثيراً منها فى العلوم الأخرى ، نظراً لأن موضوع القياس وهو الإنسان ، سلوكه ونواحى حياته العقلية والوجدانية ، هو أرقى الكائنات وأعلاها تعقيدا وأكثرها مرونة وقابلية للتغير كما أنه أقلها جميعاً قابلية للتحكم والتجريب . وبالرغم من ذلك كله فقد تقدم القياس فى علم النفس والتربية تقدماً كبيرا فى السنوات الأخيرة ؛

ومما ساعد على تقدم القياس فى علم النفس والتربية استخدامه علم الإحصاء الذى ساهم فى تطوير كثير من العلوم الأخرى . على أن المشتغلين بالقياس لا يقتصرون على استخدام ما يقدمه الإحصاء بل لقد استنبطوا لأنفسهم طرقاً إحصائية خاصة مثل طرق التحليل العاملى التى استعانوا بها فى التعرف على العوامل النفسية والعقلية . ثم شاع استخدام هذه الطرق فى كثير من العلوم الأخرى .

ويهدف هذا الكتاب إلى تزويد الطالب فى علم النفس والتربية والمشتغل بالقياس فيهما بالحقائق والنظريات الأساسية فى موضوع القياس والتقييم . إذ يجيب عن تساؤلات من مثل : متى نلجأ إلى القياس ؟ ولماذا نستخدم المقاييس ؟ ومن هو مستخدمها ؟ ومتى يحق لنا أن نستخدمها ؟ وما هى شروط إجراء الاختبار ؟ وما هى شروط إنشاء المقياس ؟ وما هى خصائص المقياس الجيد ؟ وكيف نفاضل بين أدوات القياس المختلفة ؟ وكيف نفسر نتائج القياس ؟ وما معنى النتائج والأرقام التى يوصلنا إليها المقياس ؟ وإلى أى درجة يكون الاعتماد عليها ؟ بل وكيف ننشئ برنامجاً شاملا للاختبارات فى المدرسة ؟ وكيف نفسر درجات هذه الاختبارات المقننة ودرجات التلاميذ فى الامتحانات المدرسية ؟ وما هى عيوب الطرق الحالية فى الامتحانات المدرسية ؟ وكيف نتحاشاها إيجابياً بمعنى كيف نصمم طرقاً علمية موضوعية لتقييم الطالب والمدرسة والمدرس ؟

كل هذه التساؤلات وغيرها من المشاكل التى تدور بذهن من يود تقدير الأفراد والجماعات تقديرا موضوعياً كمياً وكيفياً دقيقاً سيجد لها القارئ إجابات وحلولا . هذا وسنعرض لما حققه القياس من النتائج فى بلادنا والخارج . وبالطبع لن نقدم دليلا جامعاً لكل أدوات القياس فى بلادنا والعالم . ولكننا سنقتصر على أهمها وأكثرها دلالة متناولين مشاكل ومناهج وأهداف القياس فى التربية وقياس القدرات العقلية والمعرفية والقدرات الخاصة والاستعدادات الفارقية وقياس الشخصية ، عارضين فى كل هذا لنماذج من الاختبارات الشهيرة فى كل ناحية مفسرين وناقدين ومقيمين لطرق إنشائها واستخدامها .

ونود هنا أن ننبه إلى أن للقياس أدوات أخرى فضلا عن الاختبارات

سنعرض لأهمها ، ومنها طرق التقدير والملاحظة فى الظروف غير العادية « الاختبارات الموقفية » وطرق التقرير الذاتى وبطاقات التقويم والطرق الإسقاطية والاستفتاء كما يستخدم فى اختبارات الاتجاهات إلى آخر هذه الجوانب الهامة من الحيـــاة الإنسانية ونواحى ملاحظتها وتقدير النمو والاختلال فيها .

وسنعرض ، فيما نعرض ، للبحوث العلمية والجهود المثمرة التى تقوم بها الهيئات والأفراد حتى نعكس للقارئ صورة مصغرة طبق الأصل من مناهج معالجة المشاكل المتضمنة فى موضوعات هذا الكتاب عسى أن يزول تصور بحث مشاكلنا علمياً على أنه عمل مستحيل ، فتُحفز همم وتُدفع كفاءات وتنشط جهود من تسمح له الفرص وتتوفر له الظروف فيقوم بمعالجة ميدان استشكل مهتدياً بهذه البحوث مستعيناً بأدوات القياس كما سيعرف عنها الكثير .

وفى القسم الأول علاج أصول القياس ونظرياته ومفاهيمه الأساسية ومصطلحاته العربية والأجنبية وأدواته واختباراته مصنفة معرفة . وفى القسم الثانى تناول لكيف تنشأ الاختبارات وفقراتها وكيف يبحث صدقها وثباتها وما هى طرق اشتقاق معاييرها ثم كيف تحلل فقراتها . أما القسم الثالث فهو يعرض للقياس التربوى مشاكله ومصطلحاته والتعريف به ، وتقييم المدرس ، وتقييم التلميذ ؛ ومشكلة الامتحانات ، وبرنامج الاختبارات فى المدرسة ، وأصول ومميزات ومشاكل الاختبارات التحصيلية ، ثم إنشائها ، وتمثيل باختبارات عامة ، ثم أخرى خاصة . وينتهى القسم الثالث بعرض بحوث واختبارات محلية فى ميادين القياس التربوى .

وهكذا ينتهى بنا المجلد الأول من الكتاب

الإلمام بالقياس في التربية يجب أن ندرس القياس النفسي ، أى أن المجلد الأول ينتهى ومن هنا يبدأ المجلد الثانى الذى سيغطى كل جوانب القياس وتطبيقاته وأهمها تطبيقاته التربوية . ففى المجلد الثانى عود على بدء المجلد الأول . والله أسأل أن يكون الكتاب قد أدى ما هدفنا له .

وإلى لقاء قريب فى المجلد الثانى بإذن الله .

محمد عبد السلام أحمد

فهرس المجلد الأول

صفحة

القسم الأول

التعريف بالقياس ومفاهيمه وأدواته (١ – ١٢١)

الفصل الأول : التعريف بالقياس ... ... (٣ – ٣٦)

صفحة

الفصل الثانى : المفاهيم الأساسية فى الاختبارات النفسية والتربوية

الفصل الثالث : تضنيف وتعريف الاختبارات النفسية والتربوية وعرض للمشاكل المتعلقة بها

صفحة

القسم الثانى

بناء المقاييس النفسية والتربوية ومميزاتها (١٢٣—٣٣٥)

الصفحة

الفصل الخامس : إعداد فقرات الاختبار ... ... (١٤٧–١٧٥)

الفصل السادس : الصدق ... ... (١٧٧–٢١٥)

صفحة

الفصل السابع : الثبات ... ... (٢١٧–٢٥٢)

صفحة

الفصل الثامن : تحليل الفقرات ... ... (٢٥٣–٢٩٥)

صفحة

الفصل التاسع : المعايير ... ... (٢٩٧–٢٣٥)

الصفحة

القسم الثالث

القياس التربوى (٣٤٧ – ٧٠٠)

صفحة

الفصل الحادى عشر : تقييم المدرس ... ... (٣٦٧–٤٠٦)

الفصل الثانى عشر : تقييم التلميذ ... ... (٤٠٩–٤٤٨)

صفحة

صفحة

الفصل الثالث عشر : اختبارات المقال والاختبارات الموضوعية ... ... (٤٢٩–٤٦٤)

الفصل الرابع عشر : برنامج الاختبارات فى المدرسة (٤٦٥--٥٠٠)

الفصل الخامس عشر : خصائص ومشكلات الاختبارات التحصيلية (٥٠١--٥١٨)

صفحة

القسم الأول

التعريف بالقياس ومفاهيمه وأدواته

الفصل الأول

التعريف بالقياس

- الحاجة إلى القياس النفسى والتربوى .
- نشأة القياس فى التربية وعلم النفس .
- ماذا نقيس .
- ماهو المقياس النفسى .
- من يقيس ومتى يقيس .
- تطبيقات القياس النفسى والتربوى.
- أغراض القياس .
- تحديدات القياس .

مقـــدمة :

الإنسان إنسان بما يشترك فيه مع غيره من صفات جسمية ونفسية ومن العيش فى حضارة واحدة والتعرض لأنواع متقاربة من الخبرات العامة ويشترك مثلهم فى المجتمع الإنسانى الأكبر . فهو بذلك يشبه غيره من بنى الإنسان والإنسان فرد بما ينفرد فيه عن غيره من تكوين نفسى معين وهو بذلك يختلف عن غيره من بنى الإنسان . فمن الأمور الملحوظة ، فى الإنسان وفى سائر الكائنات الحية ، أنه رغم ما بين أفراد الجنس الواحد من تشابه وتماثل إلا أن بين أفرادها فروقاً كثيرة . وبعض هذه الفروق قد يظهر من مجرد الملاحظة العرضية العابرة ولكن كثيراً منها لا يتضح بسهولة بل يحتاج إلى التدقيق وإلى الاستعانة بوسائل مساعدة لإدراكها أو تمييزها .

على أن إدراك الفرق فى أى صفة يتطلب ، أولا ، التعرف على هذه الصفة وفهمها ثم اكتشاف واستنباط وسائل لقياس ما دق من هذه الفروق : فالشخص العادى يرى آلاف النجوم فى السماء وتبدو كلها فى نظره سواء ؛ لكنه كلما ازداد علما وخبرة بها كلما زاد إدراكا للفروق والمميزات بينها . والعالم الفلكى يحتاج إلى وسائل مساعدة لدراسة هذه الفروق كالتلسكوب ومقاييس الزمن و . . الخ . وبالمثل فمن لا يعرف صفات القطن وخصائصه لا يستطيع أن يميز بين أنواعه ورتبه المختلفة .

وقد دلت جميع الدراسات التجريبية على جماعات الحيوان أن كل جماعة من الحيوان ، مهما بدت متساوية أو متقاربة ، يوجد بين أفرادها فروق كثيرة فى النشاط التلقائى وفى درجات الحساسية وفى سرعة التعلم : فإنه ابتداء من أدنى الكائنات الحية ذات الخلية الواحدة وهى البروتوزوا وجد أن عدد المحاولات اللازمة لتكوين استجابة شرطية خاصة يتراوح فى بعضها من

٧٩ إلى ٢٨٩ فى البعض الآخر . هذا عن تجربة أجريت على ٨٢ من البروتوزوا : كما وجد فى تجربة أخرى على عدد ١١ من الأغنام أن عدد المحاولات اللازمة يترا وح بين ٣ محاولات و١٧ محاولة :

إذا كان هذا هو الحال فى المملكة الحيوانية فحرى أن تكون هذه الفروق أوسع مدى كيفاً وكمًّا فى الإنسان وهو ما أيدته بالإجماع جميع الدراسات والبحوث ونعرض لهذا فيما بعد عند الكلام عن الفروق النفسية وأنواعها :

نشأة القياس فى علم النفس والتربية :

من الطريف أن نعرف أن أول من اكتشف الفروق النفسية هو عالم فلكى . إذ حدث فى سنة ١٧٩٦ أن طرد أحد الفلكيين مساعده فى المرصد لأنه أخطأ فى رصد الزمن الذى يقضيه كوكب معين فى مروره على لوحة التلسكوب أمام هذا المساعد . وكان مقدار هذا الخطأ هو ثانية واحدة . وربما لم تكن لهذا الحادث فى ذاته أهمية لولا أن جاء العالم المشهور بسل **Bessel** سنة ١٨١٦ واهتم بجمع البيانات عن الأخطاء فى تقديرات الفلكيين المختلفين . وخلص من دراسته لهذه الفروق إلى ماعرفه بالمعادلة الشخصية **personal Equation** . ومضمونها أن الأفراد يختلفون من حيث سرعة « زمن الرجع » أى أن الزمن الذى يمر بين حدوث المثير وحدوث الاستجابة يختلف طوله من فرد إلى آخر .

وقد أدى هذا الحادث التاريخى إلى اهتمام الباحثين فى النصف الأول من القرن التاسع عشر بقياس الفروق الفردية ولم يكن هدف علماء النفس التجريبيون الأول قياس الفروق الفردية إذ كان الظن السائد أنها أخطاء ولذلك اهتموا بدراستها للتخلص منها والوصول إلى صياغة أوصاف معممة للسلوك الإنسانى أو قانون عام يصف السلوك الإنسانى . كان هذا هو الاتجاه عند ڤونت **Wundt** الذى أنشأ أول معمل لعلم النفس فى ليبزج سنة ١٨٧٩ :

كما انعكس تأثر من تعلموا على يديه بعلم وظائف الأعضاء والطبيعيات خلال بحوثهم ، فاهتموا بالحساسية للمثيرات الحسية مثل البصرية ، والسمعية ، وبدراسة زمن الرجع Reaction time . وكان لهذا الاتجاه صداه فى الاختبارات النفسية الأولى التى ركزت على الظواهر الحسية ؛

وكان لذلك الاتجاه تأثيراً آخرا ؛ إذ نبه إلى أهمية الضبط الدقيق للظروف التى تسجل فيهما الملاحظات . فمثلا ، اتضح أن ألفاظ التعليمات المعطاة للمفحوص فى تجربة زمن الرجع قد تزيد أو تقلل من سرعته فى الاستجابة وأن ، نصاعة أو لون المجال المحيط قد تغير مظهر المثير البصرى . وهكذا لزم أن توحد ظروف ملاحظة كل المفحوصين . ومن هنا أصبح تقنين الإجراء Standardization طابعا يميز الاختبارات النفسية ؛

وجاء بعد فونت تلميذه سير فرانسيس جالتون Francis Galton البيولوجى الإنجليزى الشهير (١٨٢٢ – ١٩١١) فاهتم بدراسة الوراثة عند الإنسان وتحقق من أهمية قياس مميزات الأشخاص المرتبطين وغير المرتبطين بالمفحوص . وبهذا اكتشف بدقة درجة تشابه الذرية offspring ، أو الإخوة والأخوات ، وأولاد وبنات العم أو الخال ، أو التوائم twins . وفى سنة ١٨٨٢ أنشأ معملا لعلم الإنسان القياسى anthropometry : وفيه أمكنه قياس حدة البصر والسمع ، والقوة العضلية ، وزمن الرجع ، ووظائف حسية حركية sensorimotor بسيطة أخرى : ومن ثم جمع أول وأضخم ، مجموعة من البيانات عن الفروق الفردية فى العمليات النفسية البسيطة .

وصمم بنفسه فى معمله اختبارات بسيطة طبقها ، ويطبقها البعض ، حتى الآن ، إما فى صورتها الأصلية أو المعدلة : ومن مثلها نجد ، « قضيب جالتون » للتمييز البصرى للأطوال ، « وصفارة جالتون » لتحديد أعلى مقام Pitch سمعى ، و . . الخ : وإن كان متأثرا بالفيلسوف لوك Locke ، بدليل اعتقاده بأن اختبارات التمييز الحسى تعد وسيلة لمعرفة المستوى العقلى

للشخص ، فضلا عن كتاباته الفلسفية ، إلا أنه أسدى للحركة القياسية الكثير . فقد لاحظ ، وتأكد بعده ، أن المعتوهين idiots تنقصهم فى الأغلب القدرة على تمييز الحرارة ، والبرودة ، والألم :

كما كانت لجالتون الزيادة فى تطبيق مناهج الإستبيان Questionnaire والمقياس المتدرج Rating Scale ، وكذا فى استخدام منهج « التداعى الحر free association » . وله جهد مشكور فى تطوير الطرق الإحصائية لتحليل البيانات عن الفروق الفردية بحيث يمكن أن يعالجها باحث غير مدرب رياضيا . وكذا سار من بعده تلميذه العملاق كارل پيرسون Karl Pearson .

كما أن چيمس ماكين كاتل James McKeen Cattell السيكولوچى الأمريكى له جهود كبيرة فى تطوير حركة القياس النفسى وعلم النفس التجريبى . فقد أكمل دراسة الفروق الفردية فى زمن الرجع وتعاون مع جالتون وإن لم يؤيدهما ڤونت . وأنشأ وأدار معمل علم النفس التجريبى ، ونشر حركة القياس .

واستخدم اصطلاح « الاختبار العقلى mental test » سنة ١٨٩٠ لأول مرة إذ تعرض لبطارية اختبارات لقياس الكثير من السمات الحسية الحركية والنفسية . وأشار إلى أن الوظائف العقلية يمكن أن تقاس عن طريق اختبارات التميز الحسى وزمن الرجع : ويرجع تفضيله لمثل هذه الاختبارات إلى إعانه بأن الوظائف البسيطة يمكن أن تقاس بتحديد ودقة ، بينما بدت المقاييس الموضوعية للوظائف الأكثر تعقيداً عملا لا يبشر بالخير آنئذ .

وفى نهاية القرن التاسع عشر بذلت جهود لقياس وظائف أكثر تعقيداً فى اختبارات القراءة ، والتداعى اللفظى ، والذاكرة ، والحساب البسيط . وفى سنة ١٨٩٣ دعى چاسترو Jastro إلى القيام باختبارات للعمليات الحركية ، والحسية ، والإدراكية البسيطة وإلى الاهتمام بالمعايير norms :

وكان الارتباط ضعيفاً بين التقديرات على الاختبارات المختلفة وبين المستوى العقلى الذى كان يحدده المدرسون فى تقديراتهم وفى الدرجات الأكاديمية .

المشكلات التربوية :

وفى القرن التاسع عشر أيضاً بدأ الاهتمام بتصنيف وتدريب ضعاف العقول بعد أن كانوا يعاملون بقسوة . وقد ظهرت الحاجة ماسة إلى اختبار لتشخيص وتصنيف حالات الضعف العقلى بعد أن انتشرت مؤسسات خاصة بضعاف العقول فى أوروبا وأمريكا ، خاصة وأن الأمر تطلب الوصول إلى تحديد معايير القبول والوصول إلى نظام موضوعى للتصنيف . وكان أن فرق بين المجنون insane وضعيف العقل ، فالأول يبدى اضطرابات انفعالية قد ترتبط بالتدهور العقلى وقد لا ترتبط ؛ بينما يتميز الثانى أساساً بنقص العقل الذى وضح منذ الولادة أو الطفولة المبكرة .

بدأ هذا المجهود اسكيرول Esquirol ، طبيب فرنسى ، سنة ١٨٣٨ ، بتقسيم الضعف العقلى إلى مستويات ثلاثة تبدأ بالعته idiocy . وقد أشار إلى أن المقاييس الجسمية ليست كافية وأن الاستخدام الفردى للغة هو محك موثوق به للمستوى العقلى . وعلى هذا الأساس ، ميز درجتين فى البله imbecility وثلاث درجات فى العته . ففى أعلى درجات البله ، كما قال ، يكون استخدام الكلام سهلا ميسوراً ؛ وفى الدرجة الأقل ، يكون الحديث أصعب والمفردات أقل .

ومن الجدير بالذكر أن نعلم أن المحكات الراهنة فى الضعف العقلى لغوية فى الغالب ، وأن اختبارات الذكاء اليوم مشبعة بالمحتوى اللفظى :

أما عن سيجون Seguin فهو طبيب فرنسى كان رائد تدريب ضعاف العقول ، إذ نادى ببقاء الفكرة السائدة وهى استحالة الشفاء من النقص العقلى . وفى سنة ١٨٣٧ أنشأ أول مدرسة لضعاف العقول من الأطفال ؛

وله وسائل فنية لا زالت مستخدمة فى « تدريب الحس » « وتدريب العضل » كما ساهم فى اختبارات ذكاء الأداء Performance وصمم لوحة باسمه .

وقد أعد كربلن Kraepelin ، الذى اهتم بالفحص الإكلينيكى لمرضى العقول ، سلسلة من الاختبارات لقياس ماكان يظنه عوامل أساسية فى تشخيص الفرد . استخدمت هذه الاختبارات والعمليات الحسابية لقياس أثر التدريب ، والتذكر ، والقابلية للتعب والتشتيت . كما استخدم تلميذه أوهرن Oehrn اختبارات للإدراك ، والذاكرة ، والتداعى ، والوظائف الحركية فى بحث عن العلاقات المتبادلة بين الوظائف النفسية . وتبعهم الألمانى إيبنجهوس والإيطالى فيرارى Ferrari فى الاهتمام بالاختبارات فى الحالات المرضية الأمر الذى انتقده ، فى فرنسا ، بينيه Binet وهنرى Henri قائلين إن هذه الاختبارات غالباً حركية ، ومركزة دون حق على قدرات متخصصة . وقالا إن قياس الوظائف المركبة لا يلزم أن يكون محدداً جداً ، ما دامت الفروق الفردية فى هذه الوظائف كبيرة . واقترحا قائمة من اختبارات الذاكرة ، والتخيل ، والانتباه ، والفهم ، والقابلية للإيحاء ، والفهم الحسابى ، . . . الخ . تلك هى التى أدت إلى ظهور مقاييس بينيه المعروفة .

لم يأل بينيه ومساعدوه جهداً فى استخدام كل المناهج حتى قياس السمات الجسمية ، وتحليل الخطوط ، ودراسة الكف ! ونتيجة لذلك اقتنع بأنه لا مناص من قياس الوظائف العقلية المركبة . وفى سنة ١٩٠٤ كلف بينيه وسيمون Simon من قبل وزارة المعارف الفرنسية بدراسة خطوات تعلم الأطفال دون الأسوياء . وكانا ضمن لجنة فشرعا فى مقياسهما الأول Binet-Simon Scale

وفى سنة ١٩٠٥ خرج المقياس ، مكوناً من ثلاثين بنداً item مرتبة تصاعدياً بحسب الصعوبة . وكان مستوى الصعوبة يحدد موضوعياً بتطبيق البنود على ٥٠ طفلا سوياً تتراوح أعمارهم من ٣ — ١١ سنة ، وعلى بعض

الأطفال ضعاف العقول . وكانت البنود مصممة بحيث تغطى وظائف متعددة مع التأكيد على الحكم ، والفهم ، والاستدلال ، إذ اعتبرها بينيه مكونات أساسية للذكاء . وكان الاختبار فى الأغلب لفظياً .

وفى سنة ١٩٠٨ زاد عدد البنود ، وحذف ما لم تثبت صلاحيته ، وجمعت فى مستويات السن . وعليه وضعت فى مستوى ٣ سنوات كل البنود التى يستطيع الإجابة عنها الطفل العادى الذى سنه ٣ سنوات ، وهكذا فى كل سن حتى ١٣ سنة . وهكذا أمكن حساب درجة الطفل فى الاختبار خلال « العمر العقلى mental Age » ، أى خلال عمر الطفل العادى الذى يعادله فى الأداء .

وفى سنة ١٩١١ عدل الاختبار وأصبح يقيس حتى سن الرشد . وترجم الاختبار إلى لغات كثيرة واهتم به العالم . وفى أمريكا كرس له الجهود I. M. Terman وراجعه فى جامعة ستانفورد وعرف المقياس باسم Stanford-Binet وحسبت فيه نسبة الذكاء (I. Q) intelligenece Quetient وهى النسبة بين العمر العقلى والعمر الزمنى .

نشأة القياس فى التربية :

القياس فى ميدان التربية أقدم كثيراً من القياس فى علم النفس . ولقد شعر المربون قديماً بالحاجة إلى قياس التقدم والتأخر فى تلاميذهم ، والحاجة إلى التعرف على نواحى الضعف فيهم . كما شعروا أيضاً بالحاجة إلى قياس مدى نجاح جهودهم وطرقهم فى التدريس بقياس ما يظهره تلاميذهم من تقدم فيما يدرسونه . على أن هـذا القياس كان يعتمد فى أول الأمر على الملاحظة الذاتية والآراء الشخصية ثم مر بعـد ذلك بثلاث مراحل مميزة هى :

١ ـ مرحلة الامتحانات الشفوية Oral Tests .

٢ ـ مرحلة الامتحانات التحريرية من نوع المقال essay type tests

٣ — والمرحلة الثالثة وقد بدأت حديثا ، نسبيا ، عن طريق الامتحانات الموضوعية .Objective tests أو ما يعرف «بالاختبارات الحديثة» . وسنعالج هذه الطرق فى الفصول الخاصة بالامتحانات المدرسية والاختبارات التحصيلية .

وقد ذكرنا فيما سبق أن المشكلات التربوية المتعلقة بتعليم ضعاف العقل أدت إلى دفع حركة القياس النفسى قدما إلى الأمام . ولقد كانت إجابة بعض القادة القدماء فى التربية عن السؤال « لما ذا نقيس أو نقيم ؟ » ، أننا نقيس لتمكين الطلبة اللائقين من التعليم الصالح على أيدى مدرسين أكفاء : وهذا هو الهدف الأساسى لنظم التعليم الناجحة .

ما هو المقياس ؟

يقول بين K. Bean « إن المقياس ، فى نظر التربية وعلم النفس ، هو مجموعة مرتبة من المثيرات أعدت لتقيس ، بطريقة كمية أو بطريقة كيفية ، بعض العمليات العقلية أو السمات أو الخصائص النفسية » . والمثيرات هنا قد تكون أسئلة شفوية أو أسئلة تحريرية مكتوبة وقد تكون سلسلة من الأعداد أو بعض الأشكال الهندسية أو النغمات الموسيقية أو صوراً أو رسوما . وهذه كلها مثيرات تؤثر على الفرد وتستثير استجاباته .

والاختبار يجب أن يقيس شيئاً مقصوداً كأن يعطى درجة أو قيمة أو رتبة : : الخ . وبعبارة أخرى يكون هدف الاختبار دائما أن يقيس أو أن يقيم شيئاً مقصوداً . والمقياس يجب أن يعطى نوعا من الدرجات أو أن يقدم تصنيفا وصفيا أو كليهما معاً ويعرف Cronbach كرونباخ الاختبار « بأنه طريقة منظمة لمقارنة سلوك شخصين أو أكثر » وهو يرى ضرورة هذا الشمول فى التعريف حتى لا يُظن أن الاختبار قاصر على مجموعة من الأسئلة يجيب عليها المفحوص تحريريا أو شفويا . فالاختبارات

النفسية واسعة التنوع وأصبحت أنواع المثيرات التى تستخدم لغرض الاختبار كثيرة وآخذة فى النمو بسرعة كبيرة .

وعلى ضوء هـــذا التعريف للاختبار لا تعتبر المقابلة الشخصية interviewing ، وهى أداة هامة فى دراسة الحالة ، اختباراً ، لأنها لا تقيس نوعاً واحداً من السلوك فى جميع الأشخاص . فمن النادر أن يتبع نفس الأسلوب ونفس الأسئلة فى مقابلة الأشخاص المختلفين . وهذا الاختلاف أمر يستلزمه الغرض من المقابلة وهو دراسة حالة الفرد لا مقارنته بالآخرين : على أن هناك نوعاً من المقابلة الشخصية يمكن أن تعتبر اختباراً حسب التعريف السابق وهى النوع المسمى « المقابلة الشخصية المقننة » Standardized interviews ذلك لأن هذا النوع يعتمد على أسئلة مقننة تستعمل مع جميع المفحوصين ولذلك يمكن استخدام نتائجها فى المقارنة بين الأفراد .

وعلى أساس تعريف كرونباخ للاختبار بأنه طريقة منظمة لمقارنة السلوك ، يمكن اعتبار أى جهاز أو أى وسيلة لدراسة الإحباط Frustration مثلا أو أى مجموعة من الأسئلة لدراسة الشخصية أو أى تقدير لأداء الفرد فى موقف معين من مواقف الحياة يمكن اعتبار هذه جميعاً اختبارات فهى وإن كانت فى كثير من الأحيان لا تعطينا قياساً أو تقديراً كمياً لما نختبره إلا أنها تعطى وصفاً لنواحى الشخصية أو السلوك يمكن مقارنتها فى الأفراد . والمعروف أن جميع اختبارات الشخصية تحاول وصف شخصية الفرد ككل لا أن تحلل سلوكه أو شخصيته إلى أرقام وعلاقات كمية .

والقياس فى علم النفس يقوم على ما نادى به ثورنديك Thorndike فى قوله « إذا وجد شىء ، فإنه يوجد بمقدار فإذا كان موجوداً بمقدار فإنه يمكن قياسه » .

وبالرغم من أن هذه القضية التى نادى بها ثورنديك كانت حافزاً قوياً للقياس النفسى ، إلا أنها كثيراً ما أثارت الجدل حول قيمة المقاييس التى

لا تعطى تقديرا كميا مثل كثير من اختبارات الشخصية التى تعطى تقديرات وصفية أو كيفية .

ولقد استقر الأمر أخيراً وأصبح من الأمور المسلم بها أن القياس الوصفى أو الكيفى — كما فى حالة الشخصية — ضرورى فى علم النفس حينما كان القياس الكمى متعذراً :

أما آنستازى **Anne Anastasi** فتقدم تعريفاً أوفى وأدق للاختبار النفسى فتقول « الاختبار النفسى ، أساساً ، هو مقياس موضوعى ومقنن لعينة من السلوك » والاختبارات النفسية ليست إلا عينة صغيرة ممثلة للسلوك المراد قياسه أو فحصه : فاختبار المفردات مثلا الذى يتكون من ٥٠ كلمة هو عينة من الكلمات التى يفترض أن يعرف الفرد منها قدرا معينا ، هذه الكلمات قد تقدر بالآلاف : وبالمثل فى اختبار للعمليات الحسابية أو التفكير أو . . . الخ .

وهكذا فالإخصائى النفسى يقوم بما يقوم به الكيميائى فى معمله فهو يختبر عينة صغيرة من المادة المراد اختبارها . وقد يأخذ عينة أخرى من نفس المادة ليتأكد من ثبات نتائجه أو قد يحلل عينة ثالثة من نفس المادة بطرق وأدوات أخرى كى يتأكد من صحة نتائجه وهو فى كل مرة لا يأخذ إلا عينة صغيرة جدا من المادة المراد فحصها ، وكذلك فراز القطن الذى يستخرج عينات من بالات القطن ليفحصها ليحكم على جودة ورتبة القطن كله . ولكن يجب أن يختار الفاحص هذه العينة بكل دقة بحيث تكون نموذجاً تمثل جميع المادة فى حالة الكيميائى والفراز أو تمثل جميع أنواع السلوك المراد قياسه فى حالة القياس النفسى . ويتوقف مدى تمثيل الاختبار للسلوك موضوع الفحص على عدد وطبيعة الفقرات فى العينة (أى الاختبار) .

والقيمة التشخيصية أو التنبؤية للاختبار النفسى تعتمد على المدى الذى يخدمه من ذلك الميدان الواسع من السلوك ودلالة هذا المدى . فلا يهمنا أن فردا معينا أجاب بصحة على عدد معين من المسائل الحسابية فى الاختبار

بل المهم هو مدى ارتباط هذا العدد بكل العمليات الحسابية التى يلزم أن يعرفها من نطبق عليه الاختبار .

وعلى ذلك يجب أن تؤخذ فقرات الاختبار من السلوك الذى نختبره أو نريد أن نتنبأ به ويجب أن يكون الاختبار عينة لكل السلوك المطلوب قياسه . وتختلف درجة ارتباط وتشابه فقرات الاختبار مع محتوى الميدان الواسع من السلوك المقاس من اختبار إلى آخر . فاختبار الحساب يتكون من ٢٠ مسألة مأخوذة من واقع ١٠٠ مسألة حسابية تعلمها فعلا من نجرى عليهم الاختبار . كما أن هناك اختبار للسائقين يعقد فى نفس الظروف الطبيعية التى سيعملون وسطها إذ يجرى عليهم داخل السيارة وفى طريق عام . هذان الاختباران يمثلان السلوك المراد قياسه تمثيلا يزيد فى الثانى عنه فى الأول . ومن ناحية أخرى نجد اختبارات أخرى تقل درجة تمثيلها للسلوك المراد قياسه ، على الأقل ظاهريا ، إلى حد كبير مثل اختبار الرورشاخ **Rorchach** (بقع الحبر) ، حيث نريد معرفة استجابات المفحوص تجاه غيره من الناس بتفسيرنا لاستجاباته عن بقع الحبر — مادة الاختبار . فهنا تمثيل الاختبار تمثيل رمزى ليس واقعيا .

ولكن أيا كان الاختبار ، بالرغم من الفروق الكبيرة أحياناً بين أنواع الاختبارات ، فهو عينة ممثلة للسلوك المراد قياسه والتنبؤ به . تتوقف قيمة الاختبار على مدى الارتباط الحقيقى بين أداء المفحوص له وبين أدائه فى المواقف الأخرى المماثلة من حياته الواقعية :

لماذا نقيس :

الغرض الرئيسى من القياس هو الكشف عن الفروق بأنواعها المختلفة إذ أنه لولا وجود هذه الفروق لما كانت هناك حاجة إلى القياس . وتتلخص أنواع الفروق الرئيسية فى أربعة طوائف نجملها فيما يلى :

١ ــ الفروق بين الأفراد inter-individual : ويهدف قياس هذا النوع من الفروق إلى مقارنة الفرد بغيره من أفراد فرقته الدراسية أو عمره أو بيئته فى ناحية من النواحى النفسية أو التربوية أو المهنية . لتحديد مركزه النسبى ــ فيها حتى يمكن تصنيف الأفراد إلى مستويات أو إلى جماعات متجانسة .

٢ ــ الفروق فى ــ ذات الفرد intra individual : وقياس هذا النوع من الفروق يهدف إلى مقارنة النواحى المختلفة فى الفرد نفسه لمعرفة نواحى القوة والضعف فى الفرد بالنسبة لنفسه . بمعنى مقارنة قدراته المختلفة معاً ، للتعرف على أقصى إمكانياته فى كل منها بغرض الوصول إلى تخطيط أفضل لبرامج تعليمه أو لتدريبه . كما تفيد فى توجيهه مهنياً وتربوياً حتى يحقق أكبر نجاح فى حدود إمكانياته هو :

٣ ــ الفروق بين المهن inter-Occupational : فمن المعروف أن المهن المختلفة تتطلب مستويات مختلفة من القدرات والاستعدادات والسمات . وقياس هذه الفروق يفيدنا فى الانتقاء المهنى وفى التوجيه المهنى وفى إعداد الفرد عموماً للمهن .

٤ ــ الفروق بين الجماعات inter-group : تختلف الجماعات فى خصائصها ومميزاتها المختلفة . ذلك أن الدراسات العلمية المتعددة أثبتت فروقاً بين جوانب الحياة النفسية فى كل من الجنسين ، وبين الجنسيات المختلفة ، وبين الأعمار المختلفة ، . . الخ وقياس هذه الفروق يفيدنا فى دراسة سيكولوجية الجماعات وخصائص النمو ، ودراسة العوامل التى قد تكون مسئولة عن هذه الفروق لإنماء الصالح منها والتغلب على العاطل . ولا يخفى أن قياس كل نوع من هذه الفروق ، له فائدته الكبرى فى تطوير البحوث العلمية مما يؤدى بدوره إلى تطوير المقاييس ذاتها .

والوظيفة الأساسية للاختبارات النفسية هى قياس الفروق الفردية . وإن كان تشخيص الضعف العقلى فى مبدأ الأمر هدفاً للاختبارات النفسية

وما زال كذلك فى أنواع معينـة منها . وهناك استخدامات إكلينيكية للاختبارات من مثل تشخيص الاضطراب الانفعالى ، والجناح . كما أن مشاكل التربية دفعت الاختبارات النفسية فى ماضيها دفعة قوية . وما تزال المدارس تفيد الكثير من الاختبارات فى حالات التخلف المدرسى والتفوق والتوجيه التربوى والمهنى فى التعليم الفنى والتعليم العالى .

وميدان الاختيار والتصنيف المهنيين ميدان خصب جديد حققت فيه بعض الاختبارات نجاحاً كبيراً وبخاصة تلك التى تفيد فى اختيار من يشغلون الوظائف العليا إذ أن المقابلة ليست وسيلة كافية للحكم بصلاحية طالب الوظيفة . هذا فضلا عن أن الاختبارات تكون جزءاً هاماً من إعداد الأخصائيين فى شؤون الأفراد . كذلك الحال فى اختيار وتصنيف الجنود وأمامنا الدليل ابتداء من الحرب العالمية الأولى إذ بدأت الاختبارات تغزو ميادين الخدمات العسكرية كلها .

كما أن الاختبارات النفسية تساهم الآن فى حل الكثير من المشاكل العملية ، إذ يستخدمها علم النفس الفارقى أداة فى البحث ، فى دراسة الفروق الفردية ، والتوحد بالسمات النفسية ، وقياس الفروق الجماعية ، وبحث العوامل البيولوجية والثقافية وارتباطهما بالفروق السلوكية .

وهكذا يتطلب الفهم الواعى بأى ميدان من ميادين علم النفس المعاصر أن نعرف شيئاً عن تلك الاختبارات التى لم تترك ميداناً منها إلا وحققت فيه نجاحاً . الأمر الذى يوجب الدراية ببعض الاختبارات النفسية الهامة وتقييمها وتفسير نتائجها بناء على علم بتكوينها ، ووظائفها ، وحدودها .

تطبيقات القياس النفسى والتربوى

أولا : فى التربية والتعليم :

١ – التخلف الدراسى :

تعددت الاختبارات النفسية التى تكشف عن التخلف الدراسى وأسبابه ، وهذه الاختبارات لا تعتمد على تقييم التلميذ فى المواد التى درسها فحسب بل على تقديم صورة كاملة عن جوانب الحياة النفسية للتلميذ للكشف عن العوامل التى غالباً ما تتدخل أو تؤدى إلى فشله .

٢ – فى الاختيار التربوى :

تتعدد فرص اختيار معاهد التعليم للطلبة الجدد كما تتعدد الفرص أمام التلميذ فى مراحل التعليم المختلفة لاختيار مواد دراسية معينة وترك أخرى . وهنا تزود الاختبارات المعاهد والتلميذ بصورة دقيقة عما تتطلبه دراسة معينة من تفوق فى مواد أخرى . كما تعين الاختبارات النفسية التلميذ على فهم مستواه فى مواد وقدرات معينة تؤهله لدراسات تالية أو تجعله يحجم عن متابعة دراسة معينة لأنها تتطلب ما لا يتوفر لديه .

٣ – فى التوجيه التربوى :

على نفس الأسس السابقة تفيدنا الاختبارات النفسية والتربوية فى مساعدة التلميذ على اختيار دراسة تناسبه بعد أن يكون عن نفسه صورة موضوعية متكاملة تعتمد على معرفته بقدراته واستعداداته وميوله ، وسماته .

٤ – الإرشاد التربوى :

أما فى حالة اختيار التلميذ دراسة لا تناسبه أو توجيهه إليها أو قبوله فيها

بسبب عوامل مثل درجاته المدرسية ـ ، هنا وبفضل استعدادات سابقة تمهد للاضطراب ، يكون التلميذ فى حاجة إلى حل مشكلاته المتعلقة بالدراسة والمدرسة . وهنا أيضاً تفيدنا الاختبارات فى تكشف أسباب الاضطراب وفى تبصير التلميذ بها فيعمل على تلافى ما أدى إلى اضطرابه .

٥ ــ برامج الاختبار :

لكى تقوم أجهزة التربية والتعليم بوظائفها ، ولكى يقيم رجال التربية ما حققته نظم التعليم من الأهداف العامة للتربية فى مجتمعنا ، ولكى نحسن تقدير التلاميذ ونقيم المدرسة والإدارة ، لكى نصل إلى هذا كله نحتاج إلى تخطيط وتنفيذ برامج شاملة للاختبارات النفسية والتربوية فى المدرسة . وهذه البرامج ، كما سنرى فيما بعــد ، هى تجميع ، بطريقة ما ، لبعض من الاختبارات النفسية والتربوية التى تتمشى مع الهدف المرجو من برنامج الاختبارات .

٦ ــ القراءة العلاجية :

تطبق الاختبارات النفسية والتربوية على التلاميــذ فى المدارس وفى العيادات ويقع التلميذ أحياناً فى أخطاء يعاقب عليها بالحصول على درجة منخفضة فى الاختبار . ولكن استخدام الاختبارات فى مجال القراءة العلاجية يكشف عن الأخطاء المعينة التى يقع فيها التلميذ أثناء القراءة فهو قد يخطئ دائماً فى قراءة حروف معينة أو كلمات معينة أو يخطئ فى فهم عبارة ترتبط بموقف مؤلم مرَّ به . وهنا لا نهتم بدرجة الخطأ ولكن بنوعه ونهتم أكثر من هذا بتكشف الأسباب التى يرجع إليها فشله أحياناً فى الاختبار .

٧ ــ تصنيف التلاميذ :

قد تُجرى الاختبارات فى المدرسة لغرض تقسيم التلاميذ فى الفرقة الدراسية الواحدة إلى فصول ، يكون كل فصل منها مجموعة متجانسة من

حيث المستوى العقلى أو التحصيلى أو الاقتصادى والاجتماعى حسب فلسفة المدرسة .

ثانياً فى الصناعة :

١ — اختيار العمال :

أثبتت الاختبارات النفسية والمهنية نجاحاً فى الكشف عن الأفراد الصالحين للعمل واستبعاد غير الصالحين : كما نجحت الاختبارات فى اختيار الأفراد الصالحين للتدريب وإبعاد الذين لا نتوقع تقدمهم فى التدريب ، كذلك ساعدت الاختبارات على ترشيح الأفراد الذين يستحقون الترقية أو المكافأة أو تولى المراكز الرئيسية بعد الكشف عن كفايتهم فى العمل .

٢ — تصنيف العمال :

عندما يتقدم عدد قليل من الأفراد لشغل وظائف فى مؤسسة صناعية ، تقوم المؤسسة بتصنيفهم حسب الأعمال . بعد الكشف عن قدراتهم واستعداداتهم حتى يمكن وضع كل منهم فى العمل الذى يناسبه .

٣ — التقييم :

يمكن استخدام أنواع من الاختبارات النفسية (هى اختبارات الكفاية Proficiency التى سنتحدث عنها فيما بعد) فى تقييم العمل وطرق أداء الأفراد لوظائفهم وفى تقييم الآلات. والأدوات ونظام الإشراف بل وجهاز العمل كله .

٤ — التدريب :

قلنا إن الاختبارات نجحت فى الكشف عمن يتوقع أن يفيدوا من التدريب أكثر من غيرهم . وقد نجحت الاختبارات أيضاً فى تحديد أوجه النقص فى

أداء العمال وبخاصة ما يمكن منها علاجه بتنظيم برامج التدريب . بل استطاعت الاختبارات أن تقيس مدى تقدم التدريب وتقيم كفاية المدربين القائمين بالتدريب وبرامج التدريب ذاتها .

٥ ـ الحوادث :

يكاد يكون هناك اتفاق بين علماء النفس المهنى على أن هناك أفراداً أكثر ميلا للوقوع فى الحوادث أو تسببها . وقد تمكنت الاختبارات من اكتشاف هؤلاء واستبعادهم فقل لذلك عدد الحوادث فى كثير من الأعمال .

٦ ـ الموضوعية فى الحكم :

الاختبارات الموضوعية أداة يمكن بواسطتها الحكم على الأفراد بطريقة لا تعرض القائم بالحكم للأخطاء أو للتحيز والتعصب . كما أن من نقيمهم لا يعترضون على التقييم ما دامت الأداة موضوعية مقننة . وهكذا يحجم الأفراد من المستويات المهنية المنخفضة عن التقدم للعمل إذ أن الاختبارات ستستبعدهم .

٧ ـ التوجيه المهنى :

يهدف التوجيه إلى دراسة الفرد وقدراته واستعداداته وميوله وقيمه وسماته ، ثم دراسة المهن ومقتضياتها ثم مساعدة الفرد على الالتحاق بأحد المهن . والاختبارات تفيد فى تحليل الفرد وفى تحليل العمل ، وهكذا يمكن أن نوجه الفرد إلى عمل يلائمه أكثر .

ثالثاً : فى الجيش :

كان للحربين العالميتين الفضل فى دفع حركة القياس الجمعى بالذات والاختبارات النفسية عموماً دفعة قوية . فكانت الاختبارات تطبق على

الملايين الذين أتوا من آلاف المهن ليوجهوا إلى مئات الخدمات العسكرية. وبفضل اتساع نطاق الإجزاء والإمكانيات الضخمة أمكن تحديد مقتضيات مختلف الخدمات العسكرية والأسلحة ؛ وتحليل المتقدمين ، والمواءمة بين الواجبات المطلوبة والإمكانيات المتوفرة لدى المجندين والمتطوعين ، وقد حظى سلاح الطيران بكثير من الاختبارات التى قننت فى مجاله لتطوير طرق الاختيار للطيارين والهابطين بالمظلات ولاكتشاف الأخطاء التى تتضمنها عمليات الطيران المختلفة .

رابعا : فى علم النفس الإكلينيكى :

توفر الكثير من الاختبارات على دراسة عوامل الاضطراب وتشخيصها وعلاجها ونجح فى التحليل الكيفى والكمى للتكوين النفسى للأفراد المحتاجين إلى خدمات نفسية فردية .

خامساً : فى الإدارة :

استخدام الاختبارات فى مجال الإدارة تطبيق لاستخدامها فى مجال الصناعة وإن اختلفت الأعمال فى المجالين . ولكن الخبرة التى اشتقت من مجال الصناعة فى الاختيار والتوجيه والتدريب والترفيه والتوظيف والنقل وتحديد الاختصاصات والمسئوليات ، كلها كُيفت بحسب طبيعة العمليات المتضمنة فى الإدارة وأهدافها .

كما نجحت الاختبارات ، على نفس الأسس ، عندما طبقت فى مجال الوظائف الحكومية فى نفس الوظائف تقريباً .

أغراض القياس النفسى والتربوى

أولا : المسح :

يقصد بالمسح Survey حصر الإمكانيات النفسية وتستخدم الاختبارات

النفسية والتربوية فى تحديد المستويات العقلية والوجدانية والتحصيلية لمجموعات التلاميذ والعمال والمجندين . وهذا المسح لازم لتخطيط برامج التدريب والتعليم والعلاج بعد التشخيص . فالاختبارات هنا أداة لإحصاء المستويات الموجودة وبمقارنتها بالمستويات المطلوبة ، يمكن إنشاء برامج للعلاج والتعليم والتدريب ، فى حدود الإمكانيات .

ثانيا : التنبؤ :

نقيس ونقيم الفرد والجماعة ، فى وظائف معينة ، فى وقت معين . وبافتراض ثبات السلوك الإنسانى فى حدود معينة ، ومرونته فى حدود معينة أيضاً ، وخضوعه لكل نظريات علم النفس فى حدود معينة ، كذلك يمكننا بمعرفة المستوى الحالى للفرد أن نقدر المستوى المتوقع أن يصله فى نفس الوظائف التى قسناها .

ومن تطبيقات التنبؤ يمكننا معرفة مدى الاقتصاد الهائل الذى تحققه الاختبارات فى كثير من المجالات وخاصة الصناعة . ففى دراسة لاختيار الأفراد الصالحين للعمل أدى استخدام الاختبارات إلى أن أصبحت نسبة الفاشلين فى العمل ١٠٪ وكانت ٣٠٪ من قبل . وفى دراسة لاختيار الأفراد الصالحين فى التدريب وفر أحد البنوك مبلغ ١٩٢٠٠٠ دولار .

ثالثا : التشخيص :

تستخدم الاختبارات النفسية والتربوية فى تحديد نواحى القصور وتبيان الضعف والقوة فى قدرات الفرد . وعند إجراء الاختبارات على ضوء هذا الهدف نهتم بمعرفة الجوانب التى يعانى منها الفرد أو يحتمل أن تسبب اضطرابه مستقبلا . وتعتمد الاختبارات النفسية المختلفة على طرق تشخيصية مختلفة من مثل تحليل نموذج القدرات والاستعدادات ، وتحليل الجوانب المزاجية

والانفعالية ، وتحليل تشتت الاستجابات ، كما يقدم بعض الاختبارات معادلات للتناقص أو التدهور :

رابعا : العلاج

بعد المسح والتشخيص نتعرف على نواحى الضعف وجوانب القصور ونبدأ فى تولها بدراسة أعمق لمعرفة أسبابها وديناميانها . وهنا تتكون لدينا صورة واضحة عن التكوين النفسى للفرد ، من حيث الوظائف التى قسناها . فإذا توفر لدينا كل هذا وتوفر لدينا المؤهلون للقيام بعمليات العلاج وإعادة التعليم والتكييف يسهل علينا أن نفيد من الاختبارات التى صممت بغرض العلاج أو التى يمكن استخدامها فى هذا الغرض .

تحديدات القياس :

لو رجعنا إلى المقارنة بين الكيميائى وفراز القطن ومطبق الاختبارات النفسية لو جدنا أن السيكولوچى وإن اتفق معهما فى طريقة الفحص ونوع الأداة إلا أنه يختلف عنهما فى نواحى أساسية . وهى أن تقديره ، لما يفحصه من سلوك المفحوصين ، لم يبلغ بعد دقة تقدير الكيميائى مثلا . كما وأن درجة الموضوعية التى تتوفر فى المقاييس الطبيعية لا تتوفر فى المقاييس النفسية . هذا فضلا عن أن الميزان ، الذى يزن به أخصائى القطن ، يعطى نفس الأوزان دائما لنفس الأشياء (بالات القطن) بينما يغلب ألا يكون تقدير السيكولوچى ذا ثبات مماثل . هذا إلى أن الكيميائى ، أى كيميائى ، سيعطى نفس التقدير بينما يختلف ، أحياناً ، السيكولوچيون فى تقدير سمات شخص واحد ولو اختلافا قليلا : ثم إن الكيميائى عندما يقوم بعملية معينة يعرف أكيدا أنه بصدد اختبار شق معين أو خاصية معينه بينما لا تتوفر هذه الدرجة من التأكيد فى القياس النفسى ، خاصة وأننا لا نقيس الخاصية

المعينة مباشرة إذ أننا نقيس السلوك الذى نستدل منه على هذه الخاصية .

هذا إلى أن هناك فرق جوهرى آخر ينبغى ألا يغيب عن الأذهان . وهو أننا فى حالة الاختبار الكيميائى نكون بصدد مشكلة نظرية هى فهم خاصية ما للمادة بغض النظر عن تأثير هذا الاختبار على المادة المختبرة فنحن مثلا قد نتلفها أو نحولها كلها إلى غازات ولا نبقى عليها فى صورتها الأولى . بينما فى حالة الاختبار النفسى تكون المادة المختبرة هى الفرد نفسه الذى لا يتصور أننا نحتمل أن يلحق به ضرر نتيجة للقياس ، فعلى العكس نحن نهدف أساسا إلى تحقيق مصلحته أو معاونته على تحقيق مصلحته ومصلحة جماعته أو الجنس البشرى كله .

ورغم ما حققته حركة القياس من نجاح فى ميادين القياس المختلفة وما قدمته من وسائل وأدوات بلغت درجة كبيرة من الموضوعية والدقة فى كثير من النواحى إلا أننا ما نزال ، حتى الآن ، ويبدو أننا سنظل هكذا فترة نرجو ألا تطول ، لا نزال نواجه الكثير من المشكلات الفنية فضلا عما سبق أن عرضنا له . فما يزال أمامنا العديد من الموضوعات التى لم نستطع أن نطور فيها مقاييس بلغت الموضوعية والدقة المرجوة . كما أن بعض المجالات لا يزال مبهما غامض المعالم إلى حد ما . فلم نتمكن بعد من وضع تعريفات إجرائية لكثير من سمات الشخصية ، مثلا ، بما يسهل علينا إنشاء أدوات لتقديرها . ونتيجة لهذا السبب نفسه ونتيجة لأن الإنسان ، موضوع القياس ، كائن عضوى كيميائى نفسى مركب بطريقة يستحيل معها تحليل ذلك الكل إلى جوانب بذاتها تصلح للقياس . نتيجة لكل هذا نجد أنفسنا أحياناً إزاء أدوات تقيس جوانب ليست نقية فيتشبع الاختبار ببعض ما كان يفضل ألا نقيسه ونحن بصدد جانب بذاته .

إلا أن كل هذه الاعتراضات وذلك القصور لا يجب أن يقعدنا عن

المزيد من الدراسات العلمية والعملية حتى نصل إلى قياس دقيق نقى لما نحتاج إلى التعرف عليه فقد أنجز العلم الكثير وليس ذلك على العلم بمستحيل .

من يقيس ومتى يقيس :

بعد أن تعرضنا لتحديدات الاختبارات ونواحى قصورها وما يكتنف عملية القياس النفسى والتربوى من تعقيد ، نجد أنه من اللازم أن نكون على حذر شديد عندما نستخدم الاختبارات النفسية والتربوية كأدوات للتقييم والتنبؤ والتشخيص . وتنحصر الأسباب الرئيسية التى تلزمنا بتقييد استخدام الاختبارات النفسية فيما يلى :

١ ــ تأهيل الفاحص : تختلف الشروط الواجب توافرها فيمن يقوم بإجراء الاختبار تبعا لنوع الاختبارات المستخدمة والغرض من تطبيقها . فمن الاختبارات ما يمكن أن يقوم بتطبيقه شخص على تدريب فنى متوسط مثل اختبارات التحصيل يمكن أن يجريها مدرس الفصل الذى درب عليها ، وهذا التدريب لا يستغرق فى العادة إلا فترة قصيرة ضمن البرامج العامة فى معاهد إعداد المدرسين وعلى النقيض من هذا نجد أن بعض الاختبارات النفسية مثل اختبارات الذكاء الفردية والاختبارات الإسقاطية يتحتم أن يعد لها الإخصائى النفسى إعداداً فنياً طويلا يقوم على أساس دراسة نظرية وتدريب عملى عميق تحت إشراف خبراء . هذا ويجب ملاحظة أن الطلبة الذين يجرون الاختبارات بفرض التـــدريب ليسوا أكفاء ولا يمكن الأخذ بنتائجهم أو تفسيراتهم .

على أن الحاجة إلى فاحص مؤهل تصبح ماسة فى ثلاث نواحى أساسية : الناحية الأولى اختيار القياس الذى سيستخدمه ، والناحية الثانية التطبيق والتصحيح . والناحية الثالثة هى تفسير نتائج الإجراء . هذا إلى أن الفاحص الكفء يجب أن يضع فى أعتباره تكاليف إجراء الاختبار من حيث الزمن

والجهد والمال ، كذا يجب أن ينظر إلى الاختبار من حيث سهولة الاستخدام وإمكان نقل مواده من مكان إلى آخر . ومن الاعتبارات الهامة أيضاً فى اختيار المقياس سهولة وسرعة التصحيح .

كل هذه المعلومات عادة نجدها فى كتالوجات الاختبارات ويجب أن يحسب حسابها فى تخطيط برامج الاختبار . على أن هناك مسألة أساسية فى تقييم الإختبار نفسه وهى النظر إلى ثباته وصدقه ومعاييره(*) . وهكذا يستطيع مستخدم الإختبار أن يحدد أى الاختبارات يناسب غرضاً معيناً وفئة معينة من المفحوصين .

كما يجب أن يتبع الفاحص التعليمات بدقة وأن يألف التعليمات المقننة ، وبخاصة إذا لزم أن نقارن بين نتائج تصحيح اختبار واحد أجراه فاحصون مختلفون أو عند تصحيح الاختبار بناءاً على معايير مختلفة . كما يلزم الضبط الدقيق لظروف الإجراء وإقامة علاقة طيبة بين الفاحص والمفحوص .

والتفسير الدقيق للدرجات يتطلب فهما شاملا للاختبار طبيعته ووظائفه ، وللمفحوصين طبيعتهم ومستواهم ، ولظروف الإجراء والغرض منه ، وبطبيعة الجماعات التى طبق عليها الاختبار أثناء عمليات تقنينه كما تفيدنا كثيراً المعلومات الأساسية عن المفحوص . خاصة وأن المفحوصين المختلفين قد يحصلون على درجة واحدة . وليس معنى هذا بالطبع تساويهم تماماً فى الوظائف المقاسة ، فإن تشابه النتائج لا يعنى بالضرورة تشابه الأسباب . فتفسير الدرجات المتقاربة غالباً ما يختلف . ولكى نصل إلى تفسير ملائم فى كل حالة يجب علينا أن نراعى العوامل السابقة وأن نضع فى الاعتبار أيضاً العوامل الخاصة التى قد تؤثر على درجة الفرد مثل ظروف الإجراء غير العادية ، كحدوث ضجيج قرب مكان الاختبار ، والحالة الانفعالية والجسمية

(*) انظر الفصول الخاصة بالثبات والصدق والمعايير .

الراهنة للمفحوص ، وخبراته القريبـــة وطرقه فى الإستجابة .

٢ ـ ألفة المفحوص بالإختبار : إن المغالاة فى استخدام الاختبارات النفسية والتربوية كلما تراءى للفاحص أن يقدر المفحوص تجعل المفحوصين يحصلون على تدريب وخبرة بالاختبارات النفسية . وهنا لا نستطيع أن نؤكد أن درجة الفرد تمثل مستوى قدرته بل قد تمثل هنا قوة ذاكرته القريبة أو البعيدة وأثر التدريب والمران . فالمفحوصون ، أحياناً ، يحفظون الإجابات ويستجيبون بالتذكر . وهناك خطر آخر ، فقد يعطى المدرس تلاميذه أسئلة تشبه الموجودة فى اختبار الذكاء الذى سيجرى عليهم كى يعدهم جيداً للموقف . فتزيف الاستجابات وتقل قيمة الاختبار كأداة تنبئية .

ويهمنا فى الحديث عن تأثير المران والتدريب أن نعرف مدى هذا التأثير وهل يشمل الاختبار كله أم فقرات منه أم يشمل نوعاً بأكمله من أنواع السلوك التى صمم الاختبار لقياسها . ومن الطبيعى أن خبرات التعليم التى يمر بها الفرد ، سواء كانت رسمية أم غير رسمية ، فى المدارس أم خارجها ، ستؤثر على أداء الفرد للاختبارات التى تمثل جوانب قريبة من تنظيمات السلوك التى تعلمها . ونوع التدريب ونوع العوامل التى دربت ونوع الاختبار أمور يتوقف عليها ما إذا كان المران والتدريب سيؤثران على أداء المفحوص لفقرات معينة من الاختبار المعين أم على كل الفقرات من هذا النوع المعين أم على مجال أوسع من أوجه النشاط . ولا تصبح درجة الاختبار غير صادقة بمعنى أنها لا تمثل مستوى القدرة الراهن لدى المفحوص إلا إذا كان مرور المفحوص بخبرات معينة يرفع درجته عل الاختبار دون أن يزيد فى مستوى قدرته زيادة جوهرية . أما التدريب على فقرات تماثل فقرات الاختبار ، وبخاصة فى اختبارات الذكاء ، كما حدث فى اختبار ستانفورد بينيه للذكاء مثلا ، فهو يرفع نسبة الذكاء ، ويزيد الارتفاع كلما زاد وجه الشبه بين الفقرات التى درب عليها المفحوص وفقرات الاختبار . والتكرار

أو المران عادة ما يؤثران على أداء الاختبار ، ولكن التأثير هنا أقل من الحالة السابقة . ويجب أن نلاحظ أن التدريب والمران قد يغيران من طبيعة الاختبار ما دام المفحوصون المختلفون يستخدمون مناهج مختلفة في حل المشاكل الواحدة التي يتضمنها الاختبار . إذ قد تصبح اختبارات الحساب مثلا اختبارات تذكر وسرعة نتيجـــة لأثر التدريب . وقد وجد أن الاختبارات التي يعتمد أداؤها على اكتشاف قاعدة عامة ترتفع الدرجة عليها بالتدريب إرتفاعاً يصل عند بعض المفحوصين إلى ٧٠ ٪ ويصل عند البعض الآخر إلى ٢٠٠ ٪ كما في حالة اختبارات المتاهات ورسوم المكعبات .

كما أن أداء اختبار شبيه بآخر يجعل التدريب على أحدهما يؤدى إلى زيادة الدرجة على الآخر . وكذلك إجراء صورة من الاختبار ثم إعادة الإجراء بصورة مكافئة من نفس الاختبار قد يرفع درجة المفحوص . ثم إن الشخص الذى أصبح ذا خبرة وتدريب على أغلب أنواع الاختبارات يمتاز على غيره عند أداء أى نوع من الاختبارات . فإن شيئاً شبيهاً بالاتجاه Attitude يكون قد نما لديه ، فهو مثلا لا يشعر بالغربة والحيرة فى موقف الاختبار ، كما أن ثقته تكون أعلى ، كذا فهناك تداخـــل بين الكثير من الاختبارات من حيث الوظائف ونوع المحتوى . على أن النتائج تختلف ، واختلافها لا يقتصر على درجة التشابه بين الاختبارات المختلفة ، بل يتعدى ذلك إلى السن ، والتعليم ، والخبرة السابقة للمفحوص بالاختبارات .

ماذا يجب مراعاته فى تطبيق الاختبارات :

١ — الدوافع والعلاقة الطيبة : كل اختبارات القدرة يجب أن تتوفر فيها الظروف التى تكفل للمفحوص أن يبذل أقصى جهده . ولهذا يجب أن توحد الظروف وأن يدفع كل مفحوص إلى بذل أقصى جهده فعلا ؛ وقد وجد مثلا أن الحوافز ، أيا كان مستواها ، تدفع المفحوص إلى بذل كل

جهده ، فقد ارتفعت درجات تلاميذ إحدى المدارس فى اختبارات الذكاء والحساب نتيجة استخدام التشجيع والطمأنينة .

كما قامت دراسات كثيرة على أثر الحوافز مثل الاستحسان والثناء والتشجيع ، وتدعيم مشاعر النجاح والفشل ، وتوفير ظروف المنافسه الفردية والجمعية ، ومعرفة النتائج ، وتأثير وجود ملاحظين ومعاونين فى موقف الاختبار ، وأثر الجوائز والمكافآت المالية . ولكن نتائج هذه الدراسات كانت متضاربة ، فأحيانا لم تؤد بعض الحوافز إلى تغيير جوهرى ، وأحياناً أدى بعضها إلى انخفاض درجة المفحوص . ولتوضيح ذلك يمكن أن نشير إلى التجربة التالية ، فقد اختبرت ثلاث مجموعات من طلبة المعاهد باختبارات لفظية وحسابية واختبارات شطب . وقد تم الإجراء فى ثلاث فئات من الظروف هى : (ا) ظروف مضبوطة عادية (ب) إعطاء تعليمات للمفحوص بأن يعمل بأقصى دقة ممكنة (حـ) والمجموعة الثالثة من من الظروف كان الفاحص فيها يقطع الإجراء كل ثلاث ثوان بأن يقول للمفحوصين إن المفروض أنهم أجابوا حتى الآن عن عدد معين من الفقرات . وكان هذا العدد دائما يستحيل على أغلبهم . والفئه (حـ) من الظروف تثير التوتر وتتضمن مشاعر الدونية والفشل ما دام المستوى الذى نطالب به المفحوصين يفوق إمكانياتهم ومن ثم كثرت أخطاؤهم .

كما أن عوامل الدفع تلعب دوراً كبيراً بالنسبة لنمط معين من المفحوصين أكثر مما يحدث بالنسبة لنمط آخر وبخاصة إذا كانوا مضطربين . وكذلك فإن مشاعر عدم الأمن والإحباط تؤثر على الأداء خاصة بالنسبة للأطفال . وكذلك تؤثر عوامل من مثل شعور المفحوص بأنه من جماعة أقلية عنصرية وأن الفاحص من الأغلبية التى تعادى جماعته . وهذا موجز ما نفيده من نتائج التجارب على الدوافع ؛ فهى أولا ترفع أو تخفض درجات المفحوصين وهذا الرفع أو الخفض يختلف باختلاف فئات المفحوصين . والدوافع ،

ثانيا ، يجب أن تؤخذ فى الاعتبار عند تفسير درجة المفحوص ظروف الدفع غير العادية ، خاصة إذا كان المفحوص ذا خبرات أساسية تختلف عن خبرات أفراد عينة التقنين . وثالثاً ، يجب إقامة علاقة طيبة قبل إجراء الاختبار .

ويختلف الأسلوب الفنى فى إقامة العلاقة الطيبة Rapport باختلاف نوع الاختبار ونوع المفحوصين وإعداد الفاحص . فمثلا عند اختبار أطفال ما قبل المدرسه نضع فى الاعتبار ما يتميزون به من خجل مع الغرباء ، وسرعة تشتت انتباههم ، والسلبيه والعناد . إزاء هذا يجب أن يكون الفاحص ودودا ، مرحا ، وأن يكون سلساً مطمئنا للطفل . فالطفل الخجول يلزمه وقت أطول ليألف الملابسات .

كما يجب استبعاد عامل الدهشة من جانب الفاحص . فلا يصدر منه ما ينم عن استغرابه لأداء المفحوص . وكذا يحسن أن يعلم المفحوص ، مقدما أى قبل الإجراء بمدة ، بأنه سيختبر حتى يؤهل نفسه للموقف . ويا حبذا لو وزعت كتيبات تبين طبيعة الاختبار ووظائفه وتورد مساعدات على الاحتفاظ بالهدوء والألفة بموقف الاختبار . وأن يبذل جهد لإقناع المفحوص بأنه يجب أن يهتم بأن تكون درجته صادقة ، أى أن تمثل سلوكه العقلى فى مواقف تشبه ما جاء بالاختبار فلا يهمه أن يبين ما يفضل أن يعمله ، أو ما يعتبر أنه واجب عليه . وعلى ذلك فلا داعى للغش وعدم الصراحة ، حتى لا يترتب على هذا دخول المفحوص مدرسة أو مهنة ، مثلا ، لا تناسبه وتسبب له متاعب يستطيع تجنبها إذا كان صادقاً ومتعاوناً .

ولهذا يراعى دائماً أن نبدأ الاختبار بفقرات من مستوى سهل نسبياً Preparatory Shock Absorbers كى يكتسب المفحوص الثقة بنفسه ويعتاد الإجابة عن نوع أسئلة الاختبار . هذا بالإضافة إلى إعطاء أسئلة للتمرين فى تعليمات الاختبار وأسئلة تعده للموقف .

٢ ـ الغش والتزييف :

هذه إحدى مشاكل الدفع وعلاجها إقناع المفحوص بأن صدق الدرجة (بالمعنى السابق ذكره) يجب أن يكون همه ، إذ أن تزييفها قد يسبب له المتاعب فيما بعد . وهذا الإقناع قد لا يكون جد فعال مع أنماط معينة من المفحوصين : ولكن ، على أى حال ، تؤدى ظروف الدفع السليم وإقامة علاقة طيبة إلى تقليل فرص الغش والتزييف : على أن المفحوص أحياناً ما يكون شديد الاهتمام بالوصول إلى هدف معين يتوقف على درجته فى الاختبار كأن يقبل بالمدرسة أو يختار فى المهنة ، كما أن المفحوص أحياناً ما يكون شاعراً بعدم الأمن ، وفى هاتين الحالتين يكون هناك احتمال كبير بأن يبذل محاولات لتزوير الدرجات .

وقد دلت البحوث على أن تزييف درجات الاختبار النفسى والتربوى تصنف فى الفئات الثلاث الآتية :

(ا) محاولة المفحوص أن يرفع درجتـــه فى إختبارات القدرة . (ب) محاولة الظهور بمظهر طيب فى اختبارات الشخصية . (ج) التمويه باللونية أو اللا سواء فى كل من اختبارات الشخصية واختبارات القدرة .

ويعتمد المفحوص أثناء محاولته لرفع درجته فى اختبارات القدرة على عدة وسائل مثل : (١) الاعتماد على المساعدات الخارجية . (٢) الغش من جاره فى الاختبار الجمعى . (٣) أن يأخذ زمناً أطول من المقرر بأن يبدأ قبل غيره أو أن يستمر رغم انتهاء الزمن المحدد . (٤) كما قد يعتمد المفحوص على معرفته السابقـــة لفقرات الاختبار . (٥) التخمين أو الإجابات العشوائية على ما لا يعرفه من الأسئلة كى يستفيد من عامل الصدفة (ولكن هذا يمكن تلافيه عند التصحيح بطريقة خاصة سيأتى شرحها فى فصل آخر) :

والمفحوص عند محاولته الظهور بمظهر طيب فى اختبارات الشخصية

يعتمد على أحد خصائص هذه الاختبارات ونقصد أن لكل أسئلة اختبارات الشخصية إجابة واحدة مقبولة والفرد يميل إلى أن يوافق على هذه الإجابة الطيبة ، خاصة وأن بعض الاختبارات تتطلب من المفحوص أن يجيب بنعم أو لا عن فقرة مثل : هل تشعر بالضيق عندما تقابل الغرباء ؟ . واضح أن لمثل هذا السؤال إجابة واحدة طيبة ، وهى « لا » ، والمفحوص عادة أميل إلى أن ينفى عن نفسه مثل هذا الشعور . فهو لا يجيب بما يصدق عليه فعلا بل بما يستحسنه ، أو بما يشعر أنه مقبول لدى الآخرين ، خاصة إذا كان الاختبار لغرض الانتقاء المهنى .

وقد قامت دراسة طبق فيها اختبار الشخصية لبرنرويتر Bernreuter Personality Inventory ، مرتين على مجموعة واحدة من المفحوصين . وطلب إليهم فى المرة الأولى أن يتظاهروا فى إجاباتهم بالرغبة فى الالتحاق بوظيفة بائع . وعند إعادة الإجراء طلب إليهم أن يستجيبوا مدعين الرغبة فى شغل وظيفة أمين مكتبة . كان الفاحص يطلب من المفحوصين فى مرتى الإجراء أن يجيبوا وكأنهم فعلا ، برغم رغبتهم الحقيقية ، يرغبون الالتحاق بوظيفة بائع فى المرة الأولى ، وأمين مكتبة فى المرة الثانية . ونحن نعلم أن هاتين الوظيفتين تختلفان من حيث درجة ثقة من يشغل كل وظيفة منهما بنفسه ، ورغم هذا أمكن للمفحوصين أن يجيبوا فى المرة الأولى إجابة البائع الناجح ، وفى الثانية إجابة أمين المكتبة الناجح . رغم أن الأولى تختلف عن الثانية من حيث الدرجات على مفاتيح الانبساط ، الثقة بالنفس ، الاجتماعية .

والمفحوص عند ما يود الظهور بمظهر العاجز أو غير السوى فى اختبارات الشخصية واختبارات القدرة ، يدعى مثلا أنه ناقص العقل أو أنه يعانى من اضطراب انفعالى وذلك بغرض التهرب من التجنيد أو من اختياره لعمل لا يتفق مع رغباته أو يرى فيه خطراً عليه . وقد وضعت بعض إختبارات القدرة واختبارات الشخصية حلا لهذا المشكل . إذ اعدت مفاتيح خاصة

لكشف التمويه بالدونية أو اللاسواء : والافتراض الأساسى الذى تعتمد عليه هذه المفاتيح يتلخص فى أن من يحاول التزييف يبالغ فيه . فعند تزييف اختبارات الذكاء قد ينجح المفحوص فى فقرات صعبة ثم يتعمد أن يجيب إجابة خاطئة عن فقرات أسهل ؛ مما يتنافى مع اتساق القدرة العقلية . وهنا يمكن اشتقاق المفتاح من إجابات ضعاف العقول ومقارنتها بإجابات مفحوصين طلبنا منهم أن يتصنعوا الغباء فى إجاباتهم بعد أن أعددناهم فنياً لهذا التزييف .

وأحد حلول مشكلة التزييف فى اختبارات الشخصية هو إخفاء الهدف من الاختبار عن المفحوص . فعند إجراء اختبار يقيس « الذكورة — الأنوثة » يمكن مثلا أن نقول للمفحوصين إننا نحصر الاهتمامات والاتجاهات ونقيسها . ولكن هذا الإجراء لا يتوفر فى أغلب اختبارات الشخصية إذ أن الكثير من الفقرات يكشف عن نفسه . وهناك طريقة لمعالجة الغش والتزييف ، وهى طريقة الاختيار الجبرى ، أو طريقة الأسئلة غير الشفافة . وذلك بأن تتضمن الفقرة استجابتين تبدوان متساويتين ، وعلى المفحوص أن يختار منهما واحدة تنطبق عليه . ونقصد بتساوى الاستجابتين أن المفحوص يتقبلهما بدرجة واحدة أو أن المفحوص يعتقد أن المجتمع يتقبلهما بدرجة واحدة . وهناك طرق أخرى لتلافى التزييف أوكشفه ، وسنتحدث عنها بالتفصيل فى فصل آخر . أما فى إختبارات القدرات والاختيارات التحصيلية فتوجد فرصة الإجابة بالتخمين ، ولكن لهذا علاجه باستعمال معادلة التصحيح التى سيأتى الكلام عنها فى فصل آخر .

٣ — الإعداد لموقف الاختبار :

(ا) يقوم الفاحص فى العادة بقراءة التعليمات من كراسة تعليمات الاختبار ، ولكى يكون الموقف طبيعياً ، غير متكلف ، يجب أن يألف الفاحص التعليمات اللفظية . وتأتى هذه الألفة من دراسة التعليمات وقراءتها عدة مرات قبل الإجراء حتى لا تحدث أخطاء أثناء الإجراء .

(ب) ويجب أن يقوم الفاحص مقدماً ، أى قبل الإجراء ، بإعداد مواد الاختبار فيضعها على مائدة قريبة من مائدة الاختبار ، حتى تكون فى متناوله فلا يبذل جهداً ولا يتعطل الإجراء ولا ينزعج المفحوص أو يتشتت انتباهه :

(حـ) وعند استخدام جهاز آلى فى الاختبار ، يجب الكشف مبدئياً على سلامته والتدرب على استخدامه فى سهولة .

(د) وفى الاختبار الجمعى ، يقوم الفاحص بإعداد المكان ووضع موائد الاختبار على أبعاد مناسبة وأن يتأكد من أن الموائد خالية من كل ما يعطل المفحوص أو يعينه فى الإجابة وأن يتأكد من حسن إضاءة الغرفة وتهويتها وبعدها عن مشتتات الانتباه كالضوضاء .

(هـ) ويستطيع الفاحص أن يقوم بعمل « بروفه » على مفحوص بأن يحاول معه الإجراء بالطريقة التى سيتبعها مع من سيختبرهم حقاً .

(و) يقرأ الفاحص التعليمات بصوت واضح وبسرعة معتدلة ويهتم بالتوقيت ويتأكد من أن جميع المفحوصين يتبعون التعليمات ، ويجيب عن الأسئلة الفردية للمفحوصين فى حدود ما تسمح به التعليمات ، ويمنع الغش ، ويجمع كراسات الإجابات من المفحوصين فى وقت واحد وإن لم يستطع ذلك بمفرده استعان بمن يعاونه ، ويحسن إضاءة ضوء أحمر على باب غرفة الإجراء حتى لا يدخل من يعطل الإجراء سواء كان مفحوصاً تأخر عن ميعاد الإجراء أو شخصاً آخر ، وفى الاختبارات الجمعية أيضاً يفضل أن يقف معاون للفاحص بالباب مانعاً الدخول من لحظة بدء الإجراء :

ومن المهم تحديد مدى تأثير ظروف الإجراء على درجات المفحوص ، فنشاط المفحوص يؤثر على إنتاجه خاصة لو أدى هذا النشاط إلى اضطراب انفعالى أو أى ظروف معطلة كالتعب والملل . وكذلك فالخبرة المباشرة قبل الإجراء قد تعطل أو تعرقل استجابات المفحوص .

المبادئ الخلقية فى استخدام الاختبارات النفسية والتربوية :

١ ـ يجب أن يقتصر بيع وتوزيع الاختبارات النفسية والتربوية على من يحسن استخدامها وعلى الأشخاص المؤهلين فى تطبيقها . وتختلف المؤهلات المطلوبة من اختبار إلى آخر ؛

٢ ـ يجب أن يقتصر تصحيح الاختبارات النفسية والتربوية على الأشخاص المؤهلين فى التفسير . فعندما تقدم للفرد الدرجات التى حصل عليها فى أحد الاختبارات ، يجب أن يقوم بتقديمها أخصائى مدرب فى إرشاد من يضطربون إذا أعلنوا بدرجاتهم المنخفضة ، هذا بغض النظر عن صحة أو عدم صحة الدرجات ؛

٣ ـ ويجب أيضاً ألا يسمح الناشر لنفسه بإرشاد من يشترى الاختبار إلى نوع معين من الاختبارات حتى لو كان هذا الناشر أخصائياً نفسياً أو تربوياً ؛

٤ ـ ويجب ألا ينشر الاختبار قبل ثبات صلاحيته إلا إذا كان الهدف من النشر هو إجراء دراسات وبحوث عليه . وهذا الهدف يجب أن يوضح جيداً كأن ينص فى كراسة التعليمات ، صراحة وبشكل واضح ، على أن الاختبار ما زال فى دور التجربة ؛

٥ ـ كما يجب أن يرفق بالاختبار كراسة للتعليمات تذكر فيها طريقة تطبيق الاختبار وما يقيسه وتذكر فيها المعايير ووصف دقيق لعينات التقنين وظروف التطبيق أثناء التقنين والبحوث الهامة التى أجريت على الاختبار ؛

٦ ـ يجب عدم نشر أى جزء من اختبار فى أغراض الدعاية أو الإذاعة أو الصحف لأن ذلك يؤدى إلى أضرار كثيرة أولها أن يفقد المقياس قيمته أو سريته وثانيها تكوين اتجاهات غير صحيحة نحو الاختبارات .

الفصل الثانى

المفاهيم الأساسية فى الاختبارات النفسية والتربوية

مفاهيم الإحصاء والقياس

- مقدمة .
- الدرجة الخام وتفسيرها .
- الميل المركزى .
- التشتت .
- الدرجات المحولة وأنواعها .
- الدلالة والعينات .
- الارتباط .
- التحليل العاملى .
- تحليل التباين .
- القياس والتقييم .
- تعريفات ومفاهيم عن القياس والاختبارات .

مقـــدمة :

أصبح علم النفس علما science واستقل عن الأم الكبرى : الفلسفة ، وبعد عن التصورات والمفاهيم الغامضة واقترب من العلوم التقديرية بعد أن استند إلى منهج البحث العلمى وعماده التجريب . والتجريب فى علم النفس ليس مستطاعا دائما ، خاصة فى صورته الكلاسيكية ، صورة التجارب المعملية على الحيوان التى تقيس العلاقة بين ظروف يتحكم فيها المجرب ، ونسميها المتغيرات المستقلة Independent variables وبين عامل نتوقع أن يتأثر بتغير الظروف المضبوطة نسميه المتغير المعتمد dependent variable : قرب التجريب ، بهذه الصورة ، علم النفس إلى العلوم الطبيعية فالمنهج العلمى المتبع فيهما هو هو برغم اختلاف موضوع الدراسة . ولكن هذا التجريب هنا غير ممكن دائما .

ولنا معمل آخر ، معمل كبير هو تاريخ البشرية فما لا يمكن إحداثه من تغيير فى تكوين الإنسان موضوع الدراسة هنا يمكن دراسته وتتبعه خلال ظواهر حدثت من قبل أو تحدث الآن . هذه الظواهر لا يمكن أن تدخل إلى المعمل بل نخرج نحن إليها بالإحصاء . فالإحصاء هو الأداة التى نعالج بها الظواهر لنكشف عن العلاقات بينها وما إذا كانت هذه العلاقات من نوع علاقة السبب بالنتيجة ، أم من نوع علاقة التتابع والارتباط ؛ ونحن نهتم بالعلاقات الوظيفية بين المتغيرات ، هذا النوع من العلاقات لا يمكن الجزم به من مجرد ارتباط مجموعتين من المتغيرات . ولهذا نلجأ إلى الإحصاء الذى يقدم لنا أدوات ومقاييس لا للتقدير والتقييم الكمى الموضوعى الدقيق فحسب ، بل وأيضاً يوصلنا إلى طبيعة تكوين الظاهرة موضوع الدراسة وعواملها ويمكننا فضلا عن هذا من المقارنة بينها وبين غيرها ومن اكتشاف الفروق وجوهريتها ، بل وأبعد من هذا ، يذهب

بنا الإحصاء إلى التعرف على عوامل قد لا نستخلصها دون الأساليب الفنية الإحصائية .

بل إن المنهج التجريبي بصورته الكلاسيكية كي يعطينا نتائج نستطيع أن نعممها لا بد وأن يركن إلى الإحصاء . الإحصاء إذاً أداة لا غنى عنها للوصول إلى تقديرات رقمية تمكننا من التفسير الأبعد . وهذا التفسير لا يعتمد على الإحصاء وحسب ، بل إنه مبنى أساسا على الإعداد العلمى للقائم به . وعلى هذا فالإحصاء كما رأينا أداة ، بالنسبة لنا فى هذا المجال ، تساعدنا بحسب الأرضية العلمية من وضع فروض والتحقق منها ثم تعميم نتائجها بغرض الوصول إلى قوانين أو نظريات . وإن كان الإحصاء ، بالتالى ، منهجا يساعدنا على تخطيط البحوث والتجارب ، وعلى معالجة نتائجها ، إلا أنه لا يستطيع دائما أن يقطع بالنتائج ويؤكد التعميمات . وليس هذا راجعا إلى قصور منطق الإحصاء بل إلى قصور إمكانياتنا فى قياس جوانب الحياة الإنسانية .

فالمقياس أداة موضوعية ولكنه فى الوقت نفســه عينة تمثل القدرة أو الوظيفة المراد قياسها . نحن إذاً لا نلجأ إلى كل القدرة بل إلى عنية منها ، وفى هذا ما فيه من مخاطرة . ثم إننا لا نأخذ من القدرة بل نأخذ استجابات نعتقد أنها تعكسها أو تمثلها . وهذا الاعتقاد وذلك التمثيل معرضان لكثير من الأخطاء . هذا إلى جانب صعوبات القياس الأخرى والتى سبق عرضها .

وهكذا نرى أن الإحصاء وإن كان يقدم لنا مثلا طريقة لحساب الخطأ الذى يحتمل أن نكون قد وقعنا فيه أثناء خطوات القياس المختلفة إلا أنه حتى فى هذا لا يقطع بالنتائج ، وإن كان الإحصاء قد أفاد القياس كما أوضحنا فإنه أيضاً قد أفاد منه ، فقد قدم المشتغلون بالقياس ، بفضل حاجتهم ببحوثهم ، طرقا إحصائية أصبحت أساسية ومنها التحليل العاملى .

أمر آخر نريد أن ننبه إليه ثانية وهو أن دارس القياس بل ومستخدم

الاختبار لا يلزم أن يكون إحصائياً . وإن كان مصمم الاختبار يعتمد على الإحصاء فى الكثير . إلا أننا هنا لسنا بصدد من يقومون بتصميم الاختبارات وإنما بصدد من يقومون بإجرائها . وهم فى العادة لا يحتاجون إلى فهم الأسس الرياضية للعمليات الإحصائية وكيف استخلصت المعادلات والقوانين الإحصائية ، إذ نرى أنه يكفيهم أهمية فهم الإحصاء ونواحى قصوره ، ومتى يلجئون إليه ، وكيف يستفيدون منه ، وما هى الطرق والعمليات الإحصائية التى تعينهم فى مختلف عمليات القياس على التقييم والتنبؤ . وباختصار نراهم فى حاجة إلى الإلمام بالمفاهيم الإحصائية الأساسية للقياس . فضلا عن أننا هنا لسنا ، بصدد كتاب عن الإحصاء .

الدرجة الخام وتفسيرها :

إذا قلنا إن فلاناً حصل على الدرجة ١٥ فى اختبار للمفردات . ماذا نفهم من هذا ، وكيف نفسر الدرجة ؟ الحق أنها فى ذاتها لا تعنى شيئاً ولا يمكن تفسيرها على الإطلاق . فقد تكون الدرجة النهائية فى اختبار المفردات ، هذا ، هى ١٥ فتصبح درجة فلان $\frac{١٥}{١٥}$ وهى أعلا درجة ممكنة . وقد تكون الدرجة النهائية هى ٥٠ فتصبح درجته $\frac{١٥}{٥٠}$. ولنفرض أن الدرجة النهائية هى ٢٠ أى أن درجته $\frac{١٥}{٢٠}$. و $\frac{١٥}{٢٠}$ تبدو ، ظاهرياً ، درجة مرتفعة فهى تساوى ٧٥٪ . ولكن حتى لو علمنا هذا ، هل نستطيع أن نفسر الدرجة ؟ وماذا بعد ؟

قد يكون الاختبار صعباً جداً فنعتبر ٧٥٪ درجة عالية جداً . وقد يكون الاختبار سهلا جداً فنعتبر ٧٥٪ درجة منخفضة جداً . وهكذا لا يكون للدرجة معنى إلا عندما يكون هناك مستوى نفهمها فى إطاره .

ومدرس الفصل الذى يضع مثل هذا الاختبار يضمنه بعض المفردات التى تعلمها تلاميذ فصله . أو بعض المفردات التى تقرب مما تعلموه ، أى التى تمثلها من حيث عددها ومن حيث مستوى صعوبتها . يفترض إذاً أن

الاختبار يمثل ما تعلمه التلميذ ، وإنه بالتالى ، ذو مستوى صعوبة يناسبه .

ومدرس الفصل بعد أن يصحح الاختبار يعرف أن فلاناً هذا متقدم نسبياً عن زملائه . كيف توصل إلى ذلك ؟ وضع المدرس فى ذهنه أن الاختبار يناسب قدرة المفحوص ويمثل ما تعلمه . ثم نظر إلى درجات الفصل عموماً فوجد أن نسبة كبيرة منها تقل عن درجة هذا التلميذ : أى أنه حسب متوسط درجات تلاميذ الفصل على هذا الاختبار ، أو هذا ما كان يجب عليه . ومن هنا عرف أن درجة فلان هذا عالية . فإن أراد أن يعرف مركز تلميذه بالنسبة لفصله فى هذا الاختبار لجأ إلى حساب عدد التلاميذ الذين حصلوا على درجة أقل من درجته ثم قسم عدد هؤلاء ÷ عدد تلاميذ الفصل كله فحصل على نسبتهم المئوية . ومن هنا عرف أن ١٠٪ مثلا من تلاميذ الفصل يفوقونه . أى أن رتبة فلان المئوية هى العاشرة فدرجته تزيد عن درجة ٩٠٪ من زملائه .

وإذا تصورنا درجات الفصل وأردنا أن نعرف مدى تفاوتها ومدى انسجامها : أى إذا أردنا أن نعرف درجة قرب أو بعد مختلف الأفراد عن متوسط درجة الفصل لجأنا إلى حساب التشتت .

ونحن لكى نقوم بهذه العمليات المختلفة من إيجاد النسبة المئوية والمتوسط والتشتت نحتاج إلى معرفة عدد تلاميذ الفصل والدرجة النهائية فى الاختبار والدرجة الصغرى : وفى معالجتنا لدرجات تلاميذ الفصل نحتاج إلى تنظيمها بطريقة عملية تسهل علينا هذا .

ونحن نقوم بهذه العمليات بسبب عجز الدرجة الخام عن إعطاء صورة عن مستوى الفرد فى بيان مركزه بالنسبة لمجموعته . فالدرجة الخام هى نتيجة إجراء المفحوص على الاختبار وتحسب عادة بعدد الأسئلة التى أجاب عنها إجابات صحيحة ، وقد تحسب أيضاً بأن تعطى درجة للإجابة الصحيحة عن مجموعة من الأسئلة ، وتختلف درجة مجموعة معينة من الأسئلة عن درجة

أخرى تبعاً لأهمية هذه الأسئلة فى تقدير مستوى الفرد وتبعاً لدرجة صعوبة كل مجموعة من الأسئلة . وبعد أن استخلصنا الدرجة الخام بإحدى هذه الطرق أصبح لزاماً علينا أن نقارن درجة (أ) من التلاميذ بدرجة زميله (ب) فى نفس الاختبار والأمر هنا سهل نسبياً فدرجة (ب) قد تعلو درجة (أ) أو تقل عنها . ولكن إذا أردنا أن نقارن درجة (ا) بدرجات زملائه فى الفصل اختلف الأمر . كما أن مدرس الفصل عادة يقوم بتدريس نفس المادة فى أكثر من فصل واحد ويحصل تلاميذ الفصول المختلفة على درجات مختلفة . فإذا أراد هذا المدرس أن يقارن بين درجات الفصل (س) مع بقية الفصول ، أو أن يقارن بين درجة التلميذ (ا) فى الفصل س وبين درجة التلميذ (د) فى الفصل ص اختلف الأمر أيضاً .

على أن ترتيب التلميذ فى مجموعته يكون ذا معنى وأكثر قابلية للتفسير كلما زاد عدد أفراد المجموعة التى ينتمى إليها سواء كانت فصلا أم عدداً من الفصول . وهذه الميزة تجعلنا أمام مشكلة عملية تتلخص فى طريقة تجميع درجات تلاميذ الفصل الواحد أو مجموعة الفصول بطريقة تسهل عمليات التفسير التالية .

ووسيلتنا فى هذا هى تبويب البيانات Tabulation of Data أى تجميعها فى فئات على أساس مدى قرب درجات مجموعة من الأفراد أو مدى التشابه بين الأفراد الذين تشملهم الفئة . ولكى نقوم بهذه العملية نجرى الخطوات التالية :

١ – نطرح أصغر درجة حصل عليها تلميذ من أعلى درجة وصل إليها غيره ويسمى هذا بالمدى المطلق Absolute range :

٢ – نقسم هذا المدى إلى فئات لا يزيد عددها عادة عن ٢٠ ولا يقل عن ٨ لكى نحدد طول الفئة . أو مداها أى عدد الدرجات التى تشملها .

٣ ــ نقوم بحساب عدد الأفراد الذين نالوا درجات تنحصر فى الفئة الأولى ثم الذين نالوا درجات تنحصر فى الفئة الثانية ، وهكذا ويسمى هذا بالتكرار Frequency ، ومجموع التكرارات هو مجموع عدد الأفراد ، ويكون ذلك بوضع علامة تمثل درجة الفرد أمام الفئة ، وعدد العلامات أمام الفئة الواحدة هو تكرار هذه الفئة أى عدد المرات التى وردت فيها درجات تنحصر بين حدى الفئة .

ومن الطبيعى أن يختلف عدد الأفراد أمام الفئة باختلاف الدرجة التى تحدد أول الفئة والدرجة التى تحدد آخرها ، أى تبعاً لاختلاف طول الفئة وحدود الفئات .

وقد يلزم المشتغل بالقياس وخاصة عند تجريب مقياسه بعد تأليفه أن يعرف فكرة سريعة عن الصورة التركيبية لدرجات من طبق عليهم الاختبار ، وهؤلاء قد يكونون آلافاً . ولكى يصل إلى هذا يقوم بتمثيل درجات الأفراد فى رسوم بيانية . هذا العرض البيانى graphic presentation قد يكون فى صورة مضلع تكرارى أو مدرج تكرارى أو منحنى . فتمثل الفئات على المحور الأفقى والتكرارات على المحور الرأسى وتوضع نقطة تمثل مركز كل فئة وترتفع هذه النقطة على المحور الرأسى بحسب عدد الأفراد الذين نالوا درجات تشملها هذه الفئة . وتبعد هذه النقطة على المحور الأفقى على الأقل بقدر بعد مركز هذه الفئة عن أقل درجة حصل عليها فرد . وبتوصيل هذه النقط يتكون المضلع التكرارى ومنه يمكن عمل منحنى التوزيع . وفائدة هذا تتلخص فى إعطاء فكرة سريعة وفى بيان مراكز تجمع الأفراد ومدى تفرقهم فتمثل المساحة المحددة بين العمودين عند نقطتين أو درجتين والمنحنى عدد الأفراد .

النزعة المركزية

ويقصد بالنزعة المركزية Ceatral tendency ميل أفراد المجموعة إلى أن يتجمعوا فى مركزها ، إذ نجد أن نسبة كبيرة من الأفراد يحصلون على درجات تشملها فئة . وأن نسباً متناقصة من الأفراد تحصل على درجات أقل وأن نسباً مماثلة متناقصة تحصل على درجات أعلى . والفئة أو الفئات ذات التكرار الأكبر غالباً ما تتوسط التوزيع التكرارى والرسم البيانى . ونحن نلجأ إلى قياس هذه النزعة للتجمع أو التشتت عندما نريد أن نعطى رقما يدل على درجات مجموعة من الأرقام ، فعندما نريد أن نصف المجموعة بدلا من سرد درجات أفرادها يمكننا أن نقول إن متوسط درجاتهم هو كذا . وكذا هذه تأتى عادة ، فى أبسط صورة ، من جمع درجات المجموعة كلها ثم قسمة هذا المجموع ÷ عدد الأفراد . أما فى الجداول التكرارية فيحسب المتوسط بضرب تكرار الفئة فى مركزها (مركز الفئة = $\frac{\text{مجموع بدايتها + نهايتها}}{٢}$) ثم نجمع حواصل الضرب ونقسمها على عدد الأفراد (أى مجموع التكرارات). وهناك طرق مختلفة لحساب النزعة المركزية ولكن المتوسط الحسابى (م) Mean أبسطها وأكثرها شيوعا وأفضلها دقة فى أغلب الحالات .

وإذا كان توزيع المجموعة متماثلا أى أن عدد الأفراد الذين حصلوا على درجات أعلى من المتوسط يساوى عدد من حصلوا على درجات تقل عن المتوسط ، وكان التوزيع اعتداليا بمعنى أن نسب من يزيدون عن المتوسط تتناقص تدريجيا بحيث أن أعلى الدرجات تحصل عليها نسبة صغيرة جداً من الأفراد ، وكذلك تحصل نسبة مماثلة على أصغر الدرجات ويحصل عدد معين من الأفراد على درجة تزيد عن المتوسط بمقدار ٢ مثلا وفى نفس الوقت يحصل عدد مماثل على درجات تقل بمقدار ٢ عن المتوسط وهكذا .

نقول إن كان التوزيع متماثلا واعتداليا كان المتوسط الحسابى أفضل مقياس للنزعة المركزية ، بل وكان مجموع انحرافات من يزيدون عنه (والانحرافات هنا إشارتها +) مساويا لمجموع انحرافات القيم التى تقل عن المتوسط (والانحرافات هنا إشارتها −) . أى أن المجموع الكلى للانحرافات يساوى صفراً . وهكذا نجد أنه لو كانت هناك قيم ترتفع كثيراً فى المجموعة فإنها ترفع المتوسط الحسابى لهذه المجموعة ، وبالمثل يقل المتوسط إذا كانت هناك قيم متطرفة صغيرة . ولكن هذا لا يحدث إلا عندما يكون التوزيع ملتويا أو مفرطحا وبعبارة أخرى لا يكون متماثلا اعتداليا Normal .

ويفيد المتوسط ، كما أوضحنا فى الكلام عن الدرجة الخام وتفسيرها ، فى معرفة مدى انحراف درجات أفراد نبحثهم عن مستوى المجموعة . كما أننا عند مقارنة مجموعة بأخرى نقارن بين متوسطى المجموعتين لو توفرت ظروف المقارنة الأخرى مثل إجراء نفس الاختبار على مجموعتين متماثلتين .

التشتت :

عند مقارنة أفراد المجموعة الواحدة أو مجموعتين معاً ، نستخرج قيمة واحدة لكل مجموعة ، ثم نقارن الأفراد بهذه القيمة داخل الجماعة أو نقارن القيمة التى تمثل الجماعة الأولى (المتوسط) بالقيمة التى تمثل الجماعة الثانية . ولكن هذه القيمة وحدها لا تكفى لوصف المجموعة إذ نحتاج إلى معرفة التوزيع الداخلى لدرجات الأفراد ومدى قربهم أو بعدهم عن المتوسط أى مدى تشتتهم حول المتوسط . فمقاييس التشتت إذاً هى مقاييس لقرب الأفراد من بعضهم ، أى لتجانسهم . وقد يتساوى متوسط درجات المجموعة (ا) مع متوسط المجموعة (ب) ولكن يختلف توزيع الدرجات فى (ا) عنه فى (ب) فيحصل عدد من أفراد المجموعة (ا) على درجة معينة وفى المجموعة (ب) يحصل عدد أكبر من السابق أو أصغر منه على

نفس الدرجة . وهكذا لا يفيد المتوسط وحده فى وصف المجموعة ومقارنة أفرادها ومقارنتها بمجموعات أخرى . بل يجب الاستعانه بأحد مقاييس التشتت . وسنأخذ هنا الانحراف المعيارى Standard Deviation لأنه أهم مقاييس التشتت وأكثرها استعمالا ويرمز عادة للانحراف المعيارى بالرمز ع :

ومن تعريفنا للتشتت نلاحظ أنه يعنى مدى تقارب أو تباعد الدرجات أو القيم عن المتوسط . ويعرف الانحراف المعيارى بأنه « الجذر التربيعى لمتوسط مربعات انحرافات القيم عن متوسطها الحسابى » ويعتمد فى حسابه على الخطوات التالية :

١ ــ حساب المتوسط الحسابى للقيم ويرمز له بالرمز م ؛

٢ ــ حساب انحرافات القيم وذلك بطرح المتوسط من القيم المختلفة ، أى أن الانحراف (ح) = الدرجة ــ المتوسط : (هذا فى حالة البيانات غير المبوبة أما فى البيانات المبوبة فيطرح المتوسط من مركز الفئة) ؛

٣ ــ تربيع الانحرافات (ح^{2}) :

٤ ــ تجمع مربعات الانحرافات (مجـ ح^{2}) . (فى حالة البيانات المبوبة يضرب التكرار ك فى مربع الانحراف المحسوب من مركز الفئة) :

٥ ــ يقسم مجموع المربعات هذا (مجـ ح^{2}) على عدد القيم (ن) .

٦ ــ نستخرج الجذر التربيعى لخارج القسمة السابق فيكون هو الانحراف المعيارى لهذه القيم أو الدرجات . .

∴ معادلة الانحراف المعيارى هى :

$$\text{الانحراف المعيارى ع} = \sqrt{\frac{\text{مجموع مربعات الانحرافات}}{\text{عدد الحالات}}} = \sqrt{\frac{\text{مجـ ح}^{2}}{\text{ن}}}$$

وهكذا نجد أن الانحراف المعيارى ، بما أنه يعتمد على المتوسط ، يصلح

فى حالات التوزيع الاعتدالى المتماثل : وهو يتأثر بالقيم المتطرفة ويتأثر لو استؤصلت فئة من أفراد المجموعة . والانحراف المعيارى يفيد فى معرفة طبيعة توزيع أفراد المجموعة أى مدى انسجامها وهو يتأثر بالمتوسط وبعدد المجموعة وتشتتها وبمدى الاختبار المطبق لقدرة وتدريب أفرادها . كما يفيد فى وصف المجموعة إذ يكفى ، مبدئياً ، ذكر متوسطها وانحرافها المعيارى . وبهذين المعلمين أيضاً يمكن مقارنة المجموعة بمجموعة أخرى . وأبعد من هذا فإننا بمعرفة توزيع درجات الأفراد ، نعرف مدى صعوبة الاختبار بالنسبة لهم ، فإذا كان صعباً نسبياً انخفض المتوسط وحصل عدد كبير نسبياً من الأفراد على درجات تنحرف عنه انحرافاً سلبياً (أى تقل عن المتوسط) . فإذا أضفنا فقرات أسهل ارتفع المتوسط وزاد عدد من يحصلون على درجات مرتفعة . بل ويمكن علاوة على هذا معرفة مستوى صعوبة الفقرات التى أضيفت أو التى حذفت أو التى عدلت بمقارنة توزيع الدرجات قبل وبعد الحذف أو الإضافة أو التعديل .

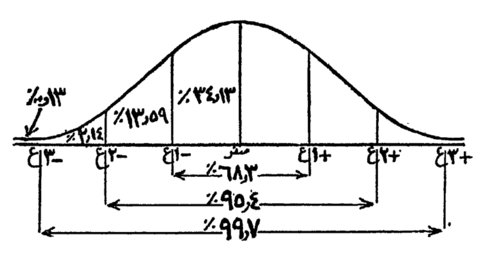
وعمليات إضافة وتعديل وتحليل واختصار فقرات الاختبار تجرى عند تحليل فقراته **item analysis** وسنتحدث بالتفصيل عن هذه الطريقة الفنية فيما بعد .

الدرجات المحولة وأنواعها

فى أول حديثنا عن الدرجة الخام وتفسيرها ذكرنا مثالا لو رجعنا إليه لوجدنا أن الدرجة التى يحصل عليها الفرد تفسر مبدئياً على أساس الدرجة النهائية للاختبار ومتوسط درجة المجموعة ، ثم عرفنا أن الانحراف المعيارى وعدد المجموعة ومستوى صعوبة الاختبار تحدد هذا التفسير . وقلنا إن أبسط طريقة هى قسمة درجة الفرد على الدرجة النهائية للاختبار ثم حساب النسبة المئوية للأفراد الذين حصلوا على درجات أعلى منها . ويأتى

ترتيبه المثوى بعد هؤلاء الأفراد . وهذا يفترض نظرياً أن المجموعة يمكن تقسيمها إلى مائة وحدة كلا منها يسمى مثين Percentile . والرتبة المثينية للفرد هى موضعه النسبي فى الجماعة الذى يحدد النسبة المئوية من الأفراد الذين حصلوا على درجات أعلى من درجته ، أو أقل منها .

وإذا أردنا أن نقارن درجة الفرد بدرجات المجموعة بغرض تفسير الأولى لزمنا أن نحولها إلى نوع من الدرجات المحولة ، والفكرة وراء هذه الدرجات المحولة هى أن تمثيل المجموعة بمنحنى يوضح مراكز أعضاء هذه المجموعة . فالمنحنى يشمل بين خطه وقاعدته كل أفراد المجموعة (تقريباً) .



(شكل ١) – المنحنى الاعتدالى مقسم إلى وحدات هى الانحرافات المعيارية .
ويوضح الشكل النسب المئوية للدرجات الواقعة داخل حدود هذه الانحرافات

فإذا كان المنحنى اعتدالياً متماثلا كان العمود الساقط من قمته على قاعدته يلتقى بها عند الدرجة المتوسطة ، وكان المماس للقمة يلتقى بالمحور الرأسى عند نقطة تمثل عدد من انحصرت درجاتهم فى فئة المتوسط . ويقسم العمود المذكور المنحنى إلى قسمين متساويين .

وإذا فرضنا أن درجة الاختبار هى ٢٠٠ وأن المتوسط الحسابى (م)

هو ١٠٠ وأن الانحراف المعيارى (ع) هو ٢٥ ، وجدنا أن م + ٣ع أى الفرد الذى حصل على درجة تزيد عن المتوسط بمقدار ثلاثة أمثال الانحراف المعيارى يحصل غالباً على أعلى درجة فى المجموعة وهى ١٠٠+٣×٢٥ = ١٧٥ وهى تقع عند ذيل المنحنى من الناحية الموجبة بينما نجد أن م — ٣ع أى ١٠٠ — ٣ × ٢٥ = ٢٥ تقع عند ذيل المنحنى من الناحية السالبة . وغالباً ما تنحصر درجات هذه المجموعة بين م + ٣ع أى (١٧٥) ، م — ٣ع أى (٢٥) ولكننا نجد أنه فى حالة التمثيل الأمثل أى الاعتــدالى ينحصر ٧٣ , ٩٩ ٪ بين م + ع ، م — ٣ع . أما إذا حصرنا عدد الأفراد الحاصلين على درجات تتراوح بين حد أعلى قدره م + ٢ ع (١٠٠ + ٢ × ٢٥=١٥٠)، وحد أدنى م — ٢ع (١٠٠ — ٢ × ٢٥ = ٥٠) فنجد نسبة هؤلاء الأفراد تعادل من مجموع الأفراد نسبة مئوية قدرها ٤٥ , ٩٥٪ وأخيراً فإننا لو حصرنا عدد الأفراد الحاصلين على م + ع كحد أعلى (١٠٠+٢٥=١٢٥) وحد أدنى م — ع (١٠٠—٢٥ = ٧٥) فنجد أنها تعادل من مجموع الأفراد نسبة مئوية قدرها ٢٧ , ٦٨٪ .

وهكذا يمكننا حساب النسبة المئوية لمن حصلوا على درجات تزيد عن المتوسط بنصف الانحراف المعيارى أو تقل عنه بعشر الانحراف المعيارى بالطريقة السابقة وهكذا . وهذه الطريقة تفيدنا فى معرفة أين يقع الفرد بالنسبة لمجموعته أى ما معنى درجته فى إطار درجات مجموعته . ونحن نلخص درجات مجموعته فى معلمين هما م ، ع وتسمى الدرجة المحولة على هذا الأساس بالدرجة المعيارية ، وتحسب بإيجاد الفرق أى الانحراف بين درجة الفرد ومتوسط مجموعته مقسوماً على الانحراف المعيارى لدرجات هذه المجموعة . وكأننا نعطى للدرجة معنى بمعرفة النسبة بين انحرافها عن المتوسط وبين انحراف درجات أفراد المجموعة عموما عن هذا المتوسط ، فإذا رمزنا

إلى درجة الفرد بالرمز س والتوسط الحسابي م والانحراف المعيارى ع فإن الدرجة المعيارية تحسب بالطريقة التالية :

$$\text{الدرجة المعيارية} = \frac{\text{س} - \text{م}}{\text{ع}}$$

وقيمة م ثابتة للاختبار الواحد وقيمة ع أيضاً ثابتة ولكن القيمتين تختلفان من اختبار إلى آخر وإن كانت أغلب الاختبارات تحاول الوصول إلى م = ١٠٠ ، ع = ١٥ وبعضها يعتمد على م = ١٠ و ع = ٣ . بينما يعتمد البعض على م = ٥ ، ع = ٢ كما فى الدرجة (ج) C. Score .

وبعض الاختبارات يستخرج متوسط درجات المجموعة التى جرب عليها وانحرافها المعيارى وتحول درجة كل فرد إلى درجة معيارية ، وتضرب هذه الدرجة المعيارية فى ١٠ ثم تجمع على + ٥٠ والنتيجة هى الدرجة التائية للفرد T score . وهنا تكون الدرجة النهائيـــة ١٠٠ ، م = ٥٠ ، ع = ١٠ . فإذا كانت درجة الفرد المعيارية هى ٢ تكون درجته التائية هى ٢ × ١٠ + ٥٠ = ٧٠ .

وهناك أنواع أخرى من الدرجات المحولة سنعرض لها تفصيلا فى فصل عن المعايير .

الدلالة

إذا قارنا بين مجموعتين فإننا نقارن بينهما على أساس المتوسط والانحراف المعيارى ، وتتأثر عملية المقارنة هذه بعدد أفراد المجموعتين : والدلالة significance تعنى الكشف عما إذا كان الفرق بين المجموعتين فرقا جوهريا برغم تطبيق اختبار واحد عليهما وتوحيد الظروف فى الحالتين . وكون الفرق جوهرياً أو ذا دلالة يعنى أن درجات أفراد المجموعتين من حيث تمركزها

وتشتها تختلف وأن الاختلاف راجع إلى طبيعة المجموعتين وليس راجعاً مثلا إلى أخطاء فى الاختبار أو فى عمليات القياس . والفرق الجوهرى إذن لا يرجع إلى الصدفة ، ويعنى أنه لا يمكن اعتبار المجموعتين أجزاءاً من مجموعة أكبر . وهناك عمليات مختلفة تجرى لقياس دلالة الفروق أهمها

« اختبار ت » T. test ومعادلته هى : $ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$

حيث $م_1$ هى متوسط الدرجات لأفراد المجموعة الأولى، $م_2$ هى متوسط الدرجات للمجموعة الثانية ، و$ع_1^2$ هى مربع الانحراف المعيارى لدرجات المجموعة الأولى ، أى التباين variance ، $ع_2^2$هى مربع الانحراف المعيارى لدرجات المجموعة الثانيـــة ، و$ن_1$ عدد أفراد المجموعة الأولى ، $ن_2$ عدد أفراد المجموعة الثانية .

وبالكشف فى جداول « اختبار ت » يمكن معرفة ما إذا كان الفرق جوهرياً أم لا ، وما مدى جوهرية هذا الفرق :

ولكى يتضح لنا كل هذا لابد أن نعرض للعينات التى نأخذها سواء من المفحوصين أو الفقرات ثم نعود إلى الدلالة مستفيدين من عرضنا للعينات فى تفسير معنى الدلالة ومستوياتها .

العينات :

لا يستطيع المشتغل بالقياس أن يطبق الاختبار ، عند تقنينه على المجموعة الأصلية Population كلها ، أى لا يستطيع أن يطبقه على كل الأفراد . إذ أن عدة عقبات عملية تحول دون ذلك وخاصة إذا كان المجتمع الأصلى المراد قياسه كبير العدد . ولذا يقوم الباحث عادة باختيار عينة من الأفراد الممثلين لهذا المجتمع الأصلى . وهو فى اختياره هذا يحاول أن يجعل كل الصفات

والخصائص المميزة للمجتمع الأصلى متوفرة فى العينة . وهو فى هذا يحافظ على تمثيل العناصر المكونة للمجتمع الأصلى تمثيلا تاماً من حيث عدد هذه العناصر . ونسبة كل عنصر منها فى المجتمع الأصلى .

فإذا كان المجتمع الأصلى مدرسة عدد تلاميذها ٣٠٠ تلميذاً ؛ ولم تتوفر إمكانيات إجراء الاختبار عليهم جميعا ، كان عليه أن يختار منهم عينة تمثلهم وليكن عدد أفرادها ١٠٠ . على أن أفراد هذه العينة يجب أن يمثلوا عدد الفصول والفرق الدراسية كأن يأخذ ٤٠ تلميذاً من السنة الأولى — ١٠ من كل فصل ، و ٣٠ تلميذاً من السنة الثانية — ١٠ من كل فصل ، و ٣٠ تلميذاً من السنة الثالثة — ١٠ من كل فصل . ذلك إن كان عدد فصول السنة الأولى ٤ وعدد فصول السنة الثانية ٣ ، وعدد فصول السنة الثالثة ٣ ، وكان عدد تلاميذ كل فصل منها هو ٣٠ تلميذاً : على أن هذه العينة وعدد أفرادها ١٠٠ يجب أن تمثل أيضاً المستويات الاقتصادية — الاجتماعية ، فإن كان عدد التلاميذ من الطبقة الاقتصادية — الاجتماعية المرتفعة هو ٦٠ تلميذاً فى المدرسة ، وجب أخذ ٢٠ منهم فى العينة ما دامت العينة هنا ⅓ المدرسة .

وبالمثل يراعى فى اختيار العينة نسبة الطبقة المتوسطة والطبقة الدنيا ، كما تراعى العوامل الأخرى التى تؤثر فيما يقيسه الاختبار المطبق مثل السن ونسبة النجاح الدراسى و . . . الخ .

والباحث فى حساب العمليات الإحصائية والنفسية يعتمد على البيانات التى اشتقها من العينة لكى يصل إلى تعميمات تنطبق على المجتمع الأصلى ؛ وعلى هذا العينة الجيدة يجب أن تتمثل فيها جميع صفات الأصل التى اشتقت منه حتى يصبح التعميم صادقاً . وعلى أى حال فيجب ألا يقل حجم العينة عن ٣٠ ؟

ولاختيار العينة طرق عدة أهمها :

أ ـ الطريقة العشوائية Ramdom : وفيها تكون الفرص متكافئة لجميع أفراد المجتمع الأصلى كى يدخلوا ضمن العينة : وهنا قد نرجع إلى كشوف التلاميذ أو العمال أو المرضى ونختار بحسب تسلسل الأرقام فإذا كانت العينة $\frac{1}{10}$ من المجتمع الأصلى اخترنا الأرقام ، ١ ، ١١ ، ٢١ . . الخ .

ب ـ الطريقة الطبقية Stratified : وقد تسمى الطريقـة الحوضية وتتلخص الفكرة فيها فى النقاط التالية :

١ ـ نصنف المجتمع الأصلى إلى فئات تبعاً للعوامل التى قد تؤثر فيما نقيسه :

٢ ـ نختار العينة العشوائية من كل فئة بحيث تمثل فئات المجتمع الأصلى كل حسب عدد أفرادها .

حـ ـ الطريقة المضبوطة Controlled : ومثال المدرسة الذى ضربناها يوضح الفكرة والخطوات المتبعة فيها .

د ـ الطريقة العرضية Accidental : وفيها يقوم الباحث باختيار أقرب الأفراد إلى إمكانياته كأن يذهب إلى مدرسة قريبة منه ويطبق الاختبار على بعض التلاميذ فى بعض الفصول دون تمثيل أو تقيد .

وهكذا نرى أن على الباحث أن يثبت صحة اختياره لعينته وذلك باختيار عينات أخرى ومقارنة النتائج .

عود إلى الدلالة

رأينا أن درجة تمثيل العينة للمجتمع الأصلى تؤثر على صحة النتائج التى نستخرجها من العينة لنعممها على الأصل . وعموماً يزيد صدق نتائجنا بزيادة عدد أفراد العينة : وتهدف الدلالة الإحصائية إلى الكشف عن مدى هذا

الصدق . ويزداد صدق التعميم كلما قربت النتائج فى العينة من النتائج فى الأصل أو فى عينة أخرى مماثلة من نفس الأصل . ويقاس هذا القرب بإيجاد الانحراف المعيارى للنتائج المستخرجة من العينة . ويسمى الانحراف المعيارى لكل معلم من هذه المعالم بالخطأ المعيارى . فمثلا الخطأ المعيارى للمتوسط هو خارج قسمة المتوسط على الانحراف المعيارى للعينة . ويتراوح المتوسط الحسابى الحقيقى ، أى مجموع درجات المجتمع الأصلى كله مقسوماً على عدد أفراده ، يتراوح بين المتوسط الذى استخرجناه من العينة مضافاً إليه مقدار الخطأ المعيارى هذا كحد أعلى ومطروحاً منه كحد أدنى وغالباً لا يزيد المتوسط الحسابى للمجتمع الأصلى كله عن (مجموع متوسط العينة + الخطأ المعيارى لهذا المتوسط) ولا يقل عن (متوسط العينة مطروحاً منه الخطأ المعيارى لهذا المتوسط) .

ولكل معلم إحصائى معادلة فى حساب دلالته .

معامل الارتباط

معامل الارتباط Coefficient of Correlation هو قيمة رياضية (عددية) تحسب بإحدى طرق عدة : وهذه القيمة العددية تبين درجة العلاقة أو التمشى بين متغيرين ، والمتغير هنا هو أى صفة أو ظاهرة نفسية تختلف من فرد لآخر أو من ظرف لآخر ويمكن قياسها أو تقديرها كمياً ، فالطول متغير والوزن متغير آخر والدرجات فى اختبار للحساب أو للذكاء متغير . . . الخ .

وتتراوح القيمة العددية لمعامل الاتباط بين + ١ ، — ١ وفيما يلى وصف لما تدل عليه بعض هذه القيم :

١ — معامل ارتباط قيمته (+ ١) : يدل هذا على علاقة تامة موجبة ، أى أن الزيادة فى أحد المتغيرين يصحبها دائماً زيادة نسبية فى التغير الثانى ،

وكذلك النقص فى أحدهما يصحبه دائماً نقص نسبى فى الثانى ، (وبعبارة أخرى يكون ترتيب الأفراد فى كلا المتغيرين ثابتاً) .

من أمثلة ذلك العلاقة بين الحجم والوزن لأجسام من نفس المادة ونفس الكثافة . أو العلاقة بين درجة الحرارة وضغط الغاز إذا كان الحجم ثابتاً . انظر الحالة ا فى شكل ٢) .

٢ — معامل ارتباط قيمته (— ١) : يدل هذا على علاقة تامة سالبة أو عكسية ، أى أن الزيادة فى أحد المتغيرين يصحبها دائماً نقص نسبى فى المتغير الثانى . أو العكس النقص فى أحد المتغيرين ، يصحبه دائماً زيادة نسبية فى المتغير الثانى ، (وبعبارة أخرى يكون ترتيب الأفراد فى أحد المتغيرين عسكه فى المتغير الثانى تماماً) ، مثل العلاقة بين حجم الغاز وضغطه إذا كانت درجة الحرارة ثابتة ، أو العلاقة بين الحجم والكثافة فى الأجسام المتساوية الوزن . (انظر الحالة هـ فى شكل ٢) .

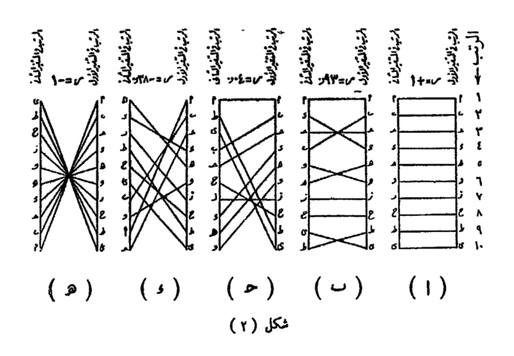
٣ — معامل ارتباط بين + ١ ، صفر أى أكبر من الصفر وأقل من + ١ صحيح . يدل هذا على علاقة جزئية موجبة أى أن الزيادة فى أحد المتغيرين كثيراً أو غالباً يصحبها زيادة نسبية فى المتغير الثانى ، وكذلك النقص فى المتغير الأول كثيراً أو غالباً ما يتبعه نقص نسبى فى المتغير الثانى . من أمثلة ذلك العلاقة بين أطوال الأفراد وأوزانهم ، أو العلاقة بين أطوال الأفراد وأعمارهم فيما تحت سن ٢٠ سنة مثلا . وكلما اقتربت قيمة معامل الارتباط من + ١ يوصف بأنه معامل ارتباط عالى موجب ومعنى ذلك أن المتغيرين يتمشيان مع بعضهما بدرجة كبيرة ، وبالعكس كلما اقتربت من الصفر دل ذلك على أن العلاقة ضعيفة . (انظر الحالة ب فى شكل ٢) .

٤ — معامل ارتباط يقع بين — ١ ، صفر أى أصغر من الصفر وأكبر من — ١ : يدل هذا على علاقة جزئية سالبة أو عكسية ، أى أن الزيادة فى أحد

المتغيرين كثيراً أو غالباً ما يصحبها نقص نسبى فى المتغير الثانى وبالعكس النقص فى المتغير الأول كثيراً أو غالباً ما يتبعه زيادة فى المتغير الثانى ومن أمثلة ذلك العلاقة بين السن وسرعة الجرى . (انظر الحالة د فى شكل ٢) ؛

٥ ــ معامل ارتباط يساوى صفر (أو صفر تقريباً) يدل هذا على عدم وجود علاقة أو ارتباط بين المتغيرين مثل العلاقة بين ذكاء الراشدين وأطوالهم . (انظر الحالة حـ فى شكل ٢) .

ويوضح الشكل (٢) رسوما تخطيطية لبعض معاملات الارتباط محسوبة بطريقة سبيرمان لحساب معامل الارتباط بالرتب Rank-Order Correlation ، وهى طريقة تعتمد على ترتيب درجات الأفراد فى كل من المتغيرين وحساب الفروق بين رتبتى كل فرد من هؤلاء الأفراد .



(أ)　(ب)　(حـ)　(ء)　(هـ)

شكل (٢)

تطبيقات على معامل الارتباط :

١ ــ معامل الثبات : هو معامل الارتباط بين نتائج الاختبار الواحد ،

ذلك بإيجاد معامل الارتباط بين مرات الإجراء المختلفة لنفس الاختبار . وقد تجرى صورة منه مرة ، ويعاد الإجراء بصورة أخرى مكافئة من نفس الاختبار . وقد يستخرج معامل الارتباط بين نصفى الاختبار الواحد . وهكذا فعامل الثبات هو ارتباط الاختبار بنفسه . وسنناقش موضوع الثبات فى فصل مستقل آخر .

٢ — معامل الصدق : هو معامل الارتباط بين نتائج الاختبار وبين ما يقيسه هذا الاختبار ، أى أن يوجد معامل الارتباط بين نتائج الاختبار وبين نتائج اختبار آخر سبق أن ثبت صدقه ، والأخير يسمى المحك Criterion . وقد يكون المحك هو مستوى النجاح فى المهنة ، أو تقديرات الخبراء ، أو النجاح المدرسى . . . الخ . وسنناقش الصدق بالتفصيل فى فصل تال .

التحليل العاملى

التحليل العاملى Factorial Analysis هو منهج إحصائى للبحث عن العوامل التى تؤثر فى الظواهر المعقدة التى ترتبط ببعضها . ويهدف التحليل العاملى إلى تفسير الظواهر الكثيرة التى يحللها إلى عدد قليل من العوامل . أى أنه يهدف إلى التبسيط العلمى .

ويعتمد التحليل العاملى على المعالجة الإحصائية لمعاملات الارتباط بين المتغيرات ، أو الاختبارات التى تقيس هذه المتغيرات ، فترصد معاملات الارتباط فى جدول متناسق الخلايا (الخانات) يسمى مصفوفة الارتباط Correlation Matrix . وفى مصفوفة الارتباط تكون خلايا العمود الرأسى الأول ، وخلايا الصف الأفقى الأول هى أرقام الاختبارات ، أى الدليل ، وترصد معاملات الارتباط فى الخلايا الداخلية وفقاً لأرقام الاختبارات فى الدليلين الرأسى والأفقى ، وتترك الخلايا القطرية خالية .

وبطريقة رياضية أو أخرى تشتق العوامل التى تلخص هذه الارتباطات أو بعضها ، ويفترض أن هذه العوامل تتعلق بسمات معينة :

ويساعد التحليل العاملى فى تنقية الاختبارات بنية الوصول إلى اختبارات نقية Pure tests يقيس كل منها سمة معينة ، كما يساعد فى حساب صدق الاختبارات ، وهو ما يعرف بالصدق العاملى Factorial validity ، وسنتاول تطبيقات طريقة التحليل العاملى فى فصل تال :

تحليل التباين

التباين Variance هو مربع الانحراف المعيارى أى أنه خارج قسمة مجموع مربعات الانحراف عن المتوسط ÷ عدد الأفراد : ونلجأ إلى تحليل التباين لمعرفة ما إذا كانت النتائج الجزئية ، التى حصلنا عليها من مصادر مختلفة متفرقة ، يمكن النظر إليها على أنها نتيجة واحدة من مصدر واحد أم لا . وأهم أداة فى هذا التحليل هى اختبار « ت » t Test ذلك لمعرفة ما إذا كان الفرق بين متوسطى مجموعتين فرقا جوهريا لا يمكن أن أن يكون قد حدث بالصدفة أى لا يمكن اعتبار العينتين من مجتمع أصلى واحد . وبالرغم من أن التباين هو مربع الانحراف المعيارى إلا أن استخدام التباين أعم فضلا عن أنه يخضع لعمليات إحصائية أخرى كثيرة : وهناك طرق أخرى لتحليل التباين لا يسمح المجال بذكرها .

القياس والتقييم

القياس Measurement و التقييم Evaluation اصطلاحان شائعان فى ميادين علم النفس والتربية ويظن البعض أنهما مترادفان وهو ظن خاطئ إذ أن ثمة فروقاً جوهرية تقوم بينهما . وأول هذه الفروق أننا نقصد بالقياس كل الملاحظات التى يمكن التعبير عنها بصورة كمية ، فنتيجة القياس مرة

نسبة الذكاء وأخرى درجة القدرة وهكذا ؛ أى أن القياس يجيب عن السؤال : ما مقدار كذا ؟ أما التقييم فهو أعم إذ أنه لا يقتصر على التحديد الكمى بل يذهب إلى أبعد من ذلك ، يذهب إلى الحكم على قيمة درجة الصفة المقاسة . فيقيم هذه الدرجة بأنها جيدة وتلك بأنها ضعيفة وهو فى هذا يهدف إلى معرفة ما إذا كانت درجة معينة من الجانب المقاس تنبئ بالنجاح أو تكشف عن حاجة إلى العلاج مثلا . أى أن التقييم يضع أهدافاً محددة يحكم على درجات الجوانب النفسية والتربوية من خلالها ، فإذا اقتربت درجة الجانب من الهدف المحدد تنبأ بالنجاح ؛

وكأن التقييم ، برغم اتصاله بالقياس واستخدامه له ، هو أعم منه لأنه يتضمن التقديرات والأحكام غير الرسمية ، الدالة مثلا على نجاح التلميذ ، وهذه الأحكام تتضمن تقدير الدرجات لمعرفة ما إذا كانت تؤهل لما يتمشى مع أهداف نرجوها ، أم لا . وعلى أى حال فإن كل قياس ناجح يجب أن يسهم فى تقدم عمليات التقييم ؛

قائمة المصطلحات
والمفاهيم الأساسية فى القياس النفسى والتربوى

اتجاه Attitude : مصطلح متعلق بأحد جوانب الشخصية يصف الميل العام العاطفى المكتسب الذى يؤثر فى الدوافع النوعية وفى سلوك الفرد وإليه يرجع السلوك المستمر المتسق نحو أو بعيداً عن مجموعة متقاربة من المواقف أو الأشياء :

أثر التدريب Practice effect : تأثير الخبرة السابقـــة المكتسبة من اختبار على الإجراء التالى لنفس الاختبار أو اختبار شبيه به . واصطلاح أثر التدريب يستخدم عندما يتدخل أو يستبعد أثر تدريب من موقف آخر غير الموقف المقاس .

اختبار الآداء Performance Test : اختبار يتطلب عادة استجابة يدوية أو استجابة حركية عموماً يقوم بها الفرد ، فى مقابل الاستجابات التى يتطلبها اختبار من نوع الورقة — والقلم Paper-and-Pencil test :

اختبار استعداد Readiness : اختبار يقيس المدى الذى حصل به الفرد درجة من النضج maturity أو اكتسب به مهارات معينة أو معلومات يتطلبها البدء فى نوع من التعلم الجديد . وهذا الاصطلاح غالباً ما يستخدم فى حالة أطفال ما قبل المدرسة لتحديد استعدادهم readiness لدخول المدرسة .

اختبار التحصيل Achievement test : مقياس للدرجة التى بها حصل الشخص أهداف التعليم أو التدريب .

الاختبار التشخيصى Diagnostic test : اختبار يستخدم فى التشخص ،

أو فى تبين نواحى القوة والضعف فى مجال معين من الدراسة . وهو يتضمن مقاييس لمركبات ، أو أجزاء متفرعة ، من جهاز أكبر من المعلومات أو المهارات .

الاختبار التنبوئى Prognostic test : اختبار يستخدم فى التنبؤ بالتحصيل المستقبل فى موضوع أو مجال معين .

اختبار التصفية Screaning test : اختبار يستخدم بغرض تصفية المتقدمين للالتحاق بعمل أو دراسة تمهيداً للقيام بعمليات أخرى لاختيارهم .

اختبار التصنيف العام General Classification test : اختبار يجرى بغرض تقسيم من يطبق عليهم إلى مجموعات تصلح كل مجموعة لمهنة أو دراسة معينة . والاصطلاح استخدم أساساً فى اختبار حربى صمم فى الحرب العالمية الثانية لتصنيف الجنود .

اختبار التمكن Mastery test اختبار غرضه الأساسى تحديد المدى الذى به تعلم أو تمكن الأفراد فى الجماعات من وحدة معينة من التدريب . وهذا النوع من الاختبارات لا يقصد منه التفريق بين الأفراد على وجه الدقة ، ولكن لتحديد ما إذا كانت أو لم تكن مجموعة من الأفراد قد حصلت مستوى معيناً من الكفاية Proficiency . وهو يستخدم لمعرفة ما إذا كانت أو لم تكن مجموعة من الأفراد مستعدة للتقدم فى وحدة تالية من التدريب .

اختبار جمعى Group test : اختبار يستطيع الفاحص أن يطبقه على عدد من الأفراد مجتمعين فى نفس الوقت .

اختبار حافل Omnibus test (وقد يسمى اختبار دشت) : اختبار يقيس فقراته عديداً من العمليات العقلية المتنوعة ، وفقراته مرتبة معاً وتعطى عنها جميعاً درجة واحدة .

اختبار ذاتى Subjective test : اختبار يصحح على أساس اتجاهات ، وتقدير ، وآراء ، وأفكار المصحح :

اختبار سرعة Speed test : اختبار له زمن محدد وفيه تعتبر سرعة الأداء أو الاستجابة أهم متغير :

اختبار فردى Individual test : اختبار يطبق على فرد واحد فقط فى المرة الواحدة .

اختبار قدرة Ability test : اختبار يتطلب أقصى أداء ليبين مستوى المركز الراهن أو القدرة الراهنة فى الوظيفة :

اختبار قوة Power test : اختبار يقصد به قياس مستوى الأداء مع تقليل أثر سرعة الاستجابة ؛ وعلى ذلك ، فهو اختبار إما لا تحديد للوقت فيه وإما أن يكون وقته متسعاً .

اختبار عبر حضارى Cross-Cultural test (وقد يسمى اختبار خالص من الثقافة Culture - free test) : وهو اختبار مصمم لاستبعاد آثار بيئة الفرد الماضية على درجته . وليس فى الإمكان عمل مثل هذا الاختبار فعلا وإن بذلت محاولات من هذا النوع ، فالاختبار الخالص من الثقافة لا يستبعد هذه الآثار ولكنه فقط بجعلها متكافئة لدى كل الأشخاص الجارى مقارنتهم :

اختبار غير لفظى Nonverbal test : هو اختبار من نوع الورقة — والقلم يستخدم مع الأطفال فى الفرق الأولى ، حيث تكون فقرات الاختبار رموزاً ، وأرقاماً وصوراً لا كلمات ؛ وتعطى التعليمات شفوياً .

اختبار لفظى Verbal test : اختبار تعتمد نتائجه بدرجة ما على استخدام وفهم الكلمات ، كما فى أغلب اختبارات الورقة — والقلم .

احتبار المسح Survey test : اختبار يقيس التحصيل العام فى موضوع

أو مجال معين . وهو عادة يعنى أكثر باتساع المدى ، منه بالتفاصيل النوعية أو اكتشاف العوامل المسببة . وهو غالباً ما يستخدم فى تصفية مجموعات كبيرة من الأشخاص .

اختبار مقنن Standardized test (وقد يسمى Objective test) : اختبار أعطى من قبل لعديد من العينات Samples أو المجموعات تحت ظروف مقننة واشتقت له معايير ؛

اختبار موضوعى Objective lest : اختبار يصحح بطريقة لا تتأثر بشخص من يقوم بتصحيحه ، فالنتائج واحدة برغم اختلاف المصححين . وتستخدم فى تصحيحه مفاتيح التصحيح أو المفاتيح المثقبة Stencils .

اختبار إجابة من عدة اختيارات Mulliple-Choice item : فقرة الاختبار التى يكون واجب الفرد فيها أن يختار الإجابة الصحيحة أو أحسن إجابة من عدد من الاختيارات المقدمة أمامه .

اختيار استجابة من عدة استجابات Multiple-responce item : نوع خاص من الفقرات التى على المفحوص فيها أن يختار استجابة واحدة أو أكثر من الاختيارات المعروضـــة أمامه ، وهى تشمل استجابتين صحيحتين أو أكثر .

الاختيار المهنى Vocational Selection : الاختيار المهنى عمليات تهدف إلى اختيار الفرد الأكثر ملاءمة ، من بين عدة أفراد متقدمين لشغل الوظيفة المعينة ، على أن ينتج أحسن إنتاج ويكون أكثر رضا .

التوجيه المهنى Vocational guidance : عمليات تهدف إلى مساعدة الفرد على أن يختار بنفسه المهنة الأكثر ملاءمة له ، من بين عدة مهن قد يمكنه الالتحاق بها ، وأن يلتحق بها ويتكيف معها ويتقدم فيها .

التوجيه التربوى Educational guidance : عمليات مساعدة التلميذ على

أن يختار بنفسه الدراسة الأكثر ملاءمة له ، من بين عدة دراسات قد يمكنه الالتحاق بها ، وأن يلتحق بها ويتكيف معها ويتقدم فيها :

الأداء المعتاد Habitual Performance : هو أداء الاختبار بحيث تمثل استجابة المفحوص سلوكه فعلا فى مواقف مشابهة لا أن يعبر الفرد عما كان يجب أن يفعل ، فى مثل هذه المواقف التى أتى بها الاختبار . وهذا النوع من الأداء نطلبه فى اختبارات الشخصية .

العامل Factor : تكوين افتراضى مستنتج من إجراء عمليات التحليل العاملى لعلاقات الترابط بين عدد من المتغيرات أو الاختبارات المتعلقة بأحد الوظائف . وهو مفهوم لوصف أحد مجالات السلوك بعد تصميم اختبار يقيس عاملا بذاته أو مجموعة من العوامل المتضمنة فى إحدى القدرات العامة للذكاء .

القدرة Ability : هى المستوى الراهن من الوظيفة سواء كان متأثراً بالتدريب أم لم يكن . والقدرة تصنيف لأجهزة من الاستجابات ترتبط فيما بينها ارتباطاً كبيراً متميزة بذلك عن غيرها من الأجهزة .

الاستعداد Aptitude : احتمال الوصول إلى درجة معينة من التمكن بعد الاستفادة من التدريب . والتدريب هنا لا يقتصر على التعليم فى المدارس والتدريب فى المصانع وإنما يشمل أيضاً الخبرات التى مر بها الفرد التى تتعلق بجانب معين من السلوك .

المهارة Skill : مستوى القيام بالعمليات الحسية الحركية المعقدة قياماً تتكامل فيه عناصر الدقة والسرعة والتكيف مع الظروف المتغيرة .

الكفاية Proficiency : درجة النجاح فى القيام بالوظائف والمقتضيات التى تتطلبها مهنة معينة :

السعة Capacity (وقد تسمى Capability) : أفضل مستوى يحتمل أن يصل إليه الفرد إذا حصل على أنسب تدريب أو تعليم .

استجابة حرة Free-response : فقرات فى فئة من الاختبارات تطلب من الفرد أن يستجيب بكلماته الخاصة لا أن يختار من بين عدة متغيرات .

استعداد مدرسى Scholastic Aptitude (وقد يسمى Academic Aptitude) : وهو الخصائص الشخصية ، الفطرية والمكتسبة ، التى تمكن من الوصول إلى درجة معينة من النجاح الأكاديمى .

الالتواء Skewness : ميل التوزيع إلى البعد عن التماثل أو الاتزان حول المتوسط .

أقصى أداء Maximum Performance : هو أداء الاختبار يبذل أقصى سرعة ودقة وقدرة لدى الفرد فى الإجابة عنه . وفى اختبار أقصى الأداء يقارن بين الأفراد على أساس أن كل فرد بذل كل ما يستطيعه إذا توفرت الظروف لذلك . وهـذا الأداء مطلوب فى كل الاختبارات ما عدا اختبارات الشخصية .

انحراف Deviation : الكمية التى بها تختلف الدرجة عن قيمة مرجعية معينة مثل المتوسط أو المعيار أو الدرجة على اختبار آخر .

الانحراف الربيعى Quartile Deviation (Q) (وقد يسمى نصف المدى الربيعى) : مقياس للتشتت يعرف على أنه نصف المسافة (بالدرجات) التى تقع بين الربيع الأعلى ر٣ (Q_3) والربيع الأدنى ر١ (Q_1) .

الانحراف المعيارى Standard Deviation (SD) : مقياس لتباين أو تشتت مجموعة الدرجات ، فكلما قربت الدرجات من المتوسط قل الانحراف المعيارى . وفى التوزيع العادى (normal) تقع حوالى ٦٨٪ من

الدرجات فى مدى ١ انحراف معيارى فوق وأقل من المتوسط ، وحوالى ٩٥٪ فى مدى ٢ انحراف معيارى . ويرمز إليه عادة بالرمز ع :

بطارية Battery : مجموعة من الاختبارات المقننة على نفس الأشخاص، ومعاييرها مشتقة بطريقة تسمح بالمقارنة : وقد يقصد بالبطارية أحياناً اختبارين أو أكثر أعطيت لنفس الأشخاص ، سواء قننت معاً أم لم تقنن :

بطارية الاستعدادات الفارقية Differental Aptitude Battery : هى مجموعة من الاختبارات قننت معاً على نفس الأشخاص بطريقة تمكن من المقارنة ، وهى تقيس استعدادات يفترق فيها الأفراد فروقاً يمكن استغلالها فى اختيارهم لمهنة أو دراسة .

التباين Variance : (انظر الانحراف المعيارى) هو متوسط مربعات الانحرافات عن المتوسط الحقيقى أو الفرضى .

تحليل الوحدات Iten analysis : عملية تقييم فقرات الاختبار الواحد باستخدام أى من الطرق المختلفة . وتتضمن هذه العملية عادة تحديد قيمة الصعوبة difficulty value والقوة التفريقية discrimination power للفقرة وعلاقاتها ببعض المحكات :

تحليل التباين Analysis of Variance : يقيس التباين الفروق الفردية والجماعية بحساب مدى انحراف كل فرد عن متوسط الأفراد ، أو مدى انحراف جماعة عن متوسط الجماعات ، أو مدى انحراف كل عينة عن الأصل الذى تنتمى إليه .

التحليل العاملى Factor Analysis : طريقة فى الإحصاء لتحديد العوامل من بين عدة ارتباطات داخلية بين مجموعة من المتغيرات ، هذه المتغيرت عادة تكون اختبارات

تصحيح أثر التخمين Correction for Gessing : تقليل الدرجة بسبب إجابات خاطئة كي تتكيف الدرجات على الاختبارات الموضوعية بحيث تلغى آثار التخمين الذى قد يقوم به الفرد فيصل إلى إجابات صحيحة . وتعتمد هذه العملية على افتراض خاطئ هو أن التخمين هو الصدفة ؛ وقد يكون من المفيد فى تقييم الدرجة أن نحدد ما إذا كانت أعلى من المستوى الذى قد توصل إليه الصدفة وحدها ، كما فى اختبار الدقة فى التنبؤ باستجابات الآخرين .

التكرار المتجمع Cumulative Frequency : هو طريقة لحساب عدد الأفراد أو الدرجات التى تزيد أو تقل عن حد معين بحيث يمكن من مجرد النظر فى عمود التكرار المتجمع ، أمام فئة معينة ، أن نعرف عدد من يقعون قبل هذه الفئة .

التقنين Standardization : يتضمن التقنين تحديد شروط تطبيق الاختبار تحديداً دقيقاً تبعاً لمبدأ مراعاة ضبط جميع العوامل التى تؤثر فى الظاهرة التى تبحث . فتوضع تعليمات الإجراء والتصحيح وتذكر المعايير ويحسن تحديد ظروف الإجراء الأخرى .

التوزيع الاعتدالى Nsrmal Distribution : توزيع الدرجات أو المقاييس بطريقة بيانية لها شكل جرسى Bell shaped : مميز وفيها تماثل حول المتوسط ، فالحالات متركزة قرب المتوسط وتتناقص فى تكرارها كلما بعدت عن المتوسط ، متمشية مع معادلة رياضية محددة .

التوزيع التكرارى Frequency Distribntion : تبويب الدرجات من الأعلى إلى الأقل أو العكس ، لتبين عدد الأفراد الذين حصلوا على كل درجة أو الذين يقعون فى كل فئة من الدرجات .

توحيد الظروف Uniformity of Conditions : وضع كل الأفراد فى

ظروف موحدة لاستبعاد أثر العوامل المتدخلة التى نرغب قياس آثارها . وتوحيد الظروف فى عمليات القياس يشمل ظروف المكان والإضاءة والتهوية والحرارة . . . الخ .

الثبات Reliability : مدى اتساق الاختبار مع نفسه فى قياس أى جانب يقيسه . أى مدى استقرار الدرجات التى يحصل عليها نفس الأفراد فى مرات الإجراء سواء أعيد الإجراء بنفس الصورة أو بصورة مكافئة من نفس الاختبار . أو كنا بصدد اتفاق درجتى الأفراد على نصفى الاختبار .

جدول الانتشار Scatter Diagram (وقد يسمى Scatter gram) : جدول ثنائى يستخدم فى توزيع الدرجات على اختبارين أو متغيرين معاً بالنسبة لكل شخص فى المجموعة . وهو يقدم دليلا بيانياً على العلاقة بين المتغيرين .

جدول الأعمار للفرق الدراسية Age-grade table : جدول يبين عدد الأفراد أو نسبهم المئوية فى كل سن داخل كل فرقة مدرسية :

اتجاه الاستجابة Response-set : ميل الفرد إلى أن يتبع نمطاً معيناً فى الاستجابة . مثل الاتجاه إلى الإجابة بـ « لا أعرف » فى اختبارات الشخصية .

الخطأ المحتمل Probable Error : دليل على تباين المقياس ، أى المدى الذى به تنحرف قيمة معينة حصل عليها الفرد عن مستواه ، وهو مقياس للخطأ فى اختيار العينة . وفى التوزيع الاعتدالى تقع نصف الانحرافات عن مقياس النزعة المركزية فى مدى خطأ محتمل واحد فوق وتحت ذلك المقياس : والخطأ المحتمل هو ٦٧٤٥ر٠ من الخطأ المعيارى .

الخطأ المعيارى Standard Error : تقدير لكمية خطأ قياس الدرجة أى الكمية التى بها تختلف درجة حصل عليها الفرد عن درجته الحقيقة التى

نفترضها . والخطأ المعيارى مأخوذ بحيث أنه فى حوالى ثلثى الحالات لا تختلف الدرجة التى حصل عليها الفرد عن درجته الحقيقية بأكثر من خطأ معيارى واحد :

الدرجة التائية T. Score ، هى درجة معيارية متوسطها ٥٠ وانحرافها المعيارى ١٠ تستخدم عادة لتحويل الدرجات الخام على اختبارين أو أكثر إلى درجات يمكن مقارنتها ، ذلك بغرض تسهيل التفسير .

الدرجة الجيمية C. Score : نوع من الدرجات المعيارية تستخدمه فئة من الاختبارات متوسطه الحسابى ٥ وانحرافه المعيارى ٢ :

الدرجة الخام Raw Score : هى النتيجة الأصلية ، التى لم تعالج إحصائياً ، التى حصلنا عليها من الاختبار أو من أى أداة قياس أخرى : وهى عادة عدد الإجابات الصحيحة ، أو النقاط التى أجاب الفرد عنها بصحة فى مقياس النقط .

الدرجة الحقيقية True Scor : الدرجة التى قد نحصل عليها إذا استخدمنا أداة قياس ثابتة تماماً . فإذا أمكن قياس الفرد عدة مرات بنفس الاختبار ، دون أن يحدث تغير فى الفرد ، قدرت درجته الحقيقية بمتوسط كل درجاته على الاختبار . ونحن دائماً لا نصل إلى درجات حقيقية بل إلى درجات نعتبرها افتراضية :

الدرجة المئينية Percentile Score : الدرجة التى تمثل النسبة المئوية من الأفراد الذين يقعون تحت درجة خام معينة .

الدرجة المعيارية Standard Score : درجة يعبر فيها عن درجة كل فرد على أساس عدد وحدات الانحراف المعيارى لدرجته عن المتوسط :

الدلالة Significance : أى مدى جوهرية النتيجة التى نحصل عليها :

وتستخدم الدلالة عادة لتبين مدى أصالة فرض وضعناه أو مدى جوهرية فرق بين متوسطين :

الربيع Quartile : أحد ثلاث نقط تقسم توزيعاً متدرجاً إلى أربعة أجزاء متساوية : والربيع الأدنى ر$_1$ (Q_1) أو المئين الخامس والعشرين ، والربيع الأعلى ر$_3$ (Q_3) أو المئين الخامس والسبعين ، والربيع المتوسط ر$_2$ (Q_2) أو المئين الخمسين وهو الوسيط Median :

العشير Decile : أحد النقط التسع التي تقسم توزيعاً مرتباً من الدرجات إلى عشر أجزاء متساوية :

السقف المنخفض Low Ceiling : اصطلاح يستخدم فى حالة الاختبار الذى يكون سهلا جداً لكثير من الأفراد ، بحيث يحصل الكثيرون على الدرجة النهائية أو أقصى درجة يسمح بها المقياس ، والعكس فى حالة السقف المرتفع (High Ceiling) إذ لا يحصل أحد على الدرجة النهائية أو أقصى درجة يسمح بها المقياس :

السمات traits : الصفة التى يمكن أن نفرق على أساسها بين فرد وآخر ويعتمد مفهوم السمة على نظرية تقول إن لكل فرد سمات شخصية ثابتة يمكن أن تلاحظ فيه . والسمات بهذا المعنى ثابتة وعامة :

السوسيوجرام Sociogram : (وقد يسمى رسم أو تخطيط العلاقات الاجتماعية) : وهو خريطة أو رسم تخطيطى يوضح التفاعلات باستخدام الأسهم فى تمثيل العلاقات والاختيارات المتبادلة بين الأفراد . وهو عادة يبين الأفراد المرغوب فيهم وغير المرغوب فيهم من أعضاء الجماعة :

السوسيومترى Sociometry (وقد يسمى مقياس العلاقات الاجتماعية) : وهو طريقة تحديد العلاقات الاجتماعية أو الشخصية بين أعضاء الجماعة .

الشخصية Personality : يقصد بهذا الاصطلاح المفهوم الذى به نصف التنظيم الدينامى الذى يتكون من تنظيمات مختلفة متكاملة تضم أجهزة السلوك والإدراك التى تعتمد على التفاعل النفسى الجسمى . على أن هذا التنظيم يحدد الطريقة الفريدة التى يتميز بها الشخص عن غيره فى استجاباته للبيئة ، عالم الأشخاص ، وعالم الأشياء ، ونفسه ، مع التأكيد على مدى التوافق فى المجال الاجتماعى .

الصدق Validity : هو أن يقيس الاختبار ما افترض أنه يقيسه . ويعرف الصدق على أساس أغراض مختلفة أى أن أنواعاً مختلفة من الأدلة تستخدم فى تحـديد أنواع الصدق وأكثر أنواع الصدق شيوعاً هى : صدق المحتوى Content validity ، التى تصف مدى جودة تمثيل محتوى الاختبار لفئة من المواقف أو الموضوعات التى نقيسها .

الصدق التلازمى Concurrent validity : الذى يصف مدى ارتباط الدرجة على الاختبار بمحكات الأداء الراهنة أو مركز الفرد حالياً ، والصدق التنبؤى Predective validity : ويدل على مدى مسايرة التنبؤات التى نصل إليها من الاختبار مع دليل نجمعه فى وقت تال على إجراء الاختبار ، وصدق البناء construct validity الذى يدل على الدرجة التى بها تعد تكوينات تفسيرية أو مفاهيم معينة ، تعد مسئولة عن الأداء فى الاختبار .

الصدق الظاهرى Face valieity (قد يسمى الصدق السطحى) : الصدق الظاهر للاختبار هو كيف يبدو الاختبار مناسباً وملائماً للفرد الذى نقيسه . والمدى الذى به يتكون الاختبار من فقرات تبدو مرتبطة بالمتغير الذى يقاس .

الصفحة النفسية Profile : رسم بيانى لنتائج أداء الفرد لمجموعة من الاختبارات أو الاختبارات الفرعية .

صلب السؤال Stem : الجزء المقدم من سؤال يختار فيه الفرد أحداً من عدة اختيارات ، وهذا الجزء على الفرد أن يكمله باختيار أحد الاستجابات التى تتبع السؤال .

الصورة المصغرة للعمل Job Miniature : طريقة من طرق صياغة فقرات الاختبار بحيث يكون التشابه تام بين عينة السلوك الذى يقيسه الاختبار وبين السلوك المراد التنبؤ به .

الصورة المكافئة Equivalent Form (وقد تسمى الصورة الموازية Parallel form أو الصـــورة المساوية Equal form) : أحدى صورتين أو أكثر من الاختبار الواحد التى تتوازى كثيراً جداً فى مضمون الفقرات وصعوبتها ، والتى تؤدى إلى متوسطات للدرجات قريبـــة جداً فى الصور المختلفة ، والتى تقيس الاختلاف وتقدر الثبات فى مجموعة معينة من الأفراد .

طابع القدرات للمهنة Occupational ability Patterns (وقد يسمى نموذج القدرات المهنية) : والمقصود به نمط تنظيم مجموعة القدرات التى يقتضيها النجاح فى مهنة معينة . ويتم ذلك بتحديد متوسطات الدرجات لقدرات الناجحين فى هذه المهنة . وتعد طريقة تساعد فى عمليات الاختيار والتوجيه المهنى والتربوى ، فمن يقترب الرسم البيانى ، الذى يمثل قدراته ، من هذا الطابع يختار لهذه المهنة أو الدراسة أو يوجه إليها ، وبالمثل يحدد طابع الاستعدادات للمهنة وكذلك طابع الميول للمهنة .

الطريقة الإسقاطية Projective Technique : إحدى طرق الدراسة فى الشخصية وفيها يكون الفرد حرا فى استجاباته بسلسلة من المنبهات مثل بقع الحبر والصور والجمل الناقصة . . . الخ . وفى موقف الاستجابة ــ الحرة هذا يسقط المفحوص فى استجاباته جوانب شخصيته المميزة .

العلاقة المستقيمة Linear Relationship : هى العلاقة بين متغيرين أو أكثر والتى يمكن تمثيلها بخط مستقيم ، وعلى ذلك فإن تزايد أو تناقص أحد المتغيرين يرتبط بزيادة أو نقص الآخر . بنسبة ثابتة على طول التغير . سواء كانت زيادة المتغير الأول تتبعها زيادة مماثلة فى الآخر أم نقصاً بنفس المعدل .

العلاقة المنحنية : Curvilinear Relationship : علاقة بين كميتين متغيرتين يمثلها منحنى فزيادة المتغير الأول تؤدى إلى زيادة المتغير الثانى . أو نقصه ولكن بمعدل غير ثابت :

عدم التجانس Hetrogeneity : ميل الجماعة إلى أن تكون غير متماثلة .

العمر الزمنى (C.A.) Chronological Age : الفترة منذ ميلاد الفرد حتى تاريخ إجراء الاختبار . وتلجأ بعض الاختبارات إلى تعديل هذه الفترة بضربها فى معامل خاص وفقاً لظروف خاصة بالاختبار ويرمز إليه فى الكتب العربية (ع ز) .

العمر التحصيلى Achievement Age : هو العمر الذى تكون درجة الاختبار التحصيلى تقديراً له أو تقديراً متوسطاً له . وقد يسمى العمر التعليمى Educational Age :

العمر العقلى (M.A.) Mental Age : العمر الذى تكون الدرجة فى اختبار الذكاء أو القدرة المدرسية متوسطاً له . والعمر العقلى هو متوسط العمر للأفراد الذين يحصلون على الدرجة المتوسطة . ويرمز إليه فى الكتب العربية بالرمز (ع ع) .

العمر القرائى Reading Age : وهو الدرجة المكافئة للعمر التى تعطى للدرجة المتوسطة على اختبار القراءة المطبق على أفراد من عمر زمنى معين .

عينة عشوائية Rondom Sample : عينة مأخوذة بحيث أن كل فرد

من المجموعة الأصلية Population أعطيت له فرصة مكافئة لأن يؤخذ ضمن العينة .

عينة ممثلة Representative Sample : هى العينة التى تكافئ المجموعة الأصلية من حيث مكوناتها ومستوياتها ونسب كل فئة فيها .

فقرة الاسترجاع Recall Item : وهى فقرة الاختبار التى تتطلب من الشخص أن يقدم الإجابة الصحيحة من ذاكرته .

فقرة التعرف Recognition Item : فقرة الاختبار التى تتطلب من الفرد أن يختار أو يتعرف على الإجابة الصحيحة من بين اختيارين أو أكثر قدما له .

فقرة التكملة Completion Item : سؤال فى الاختبار يتطلب التكملة أو ملء الفراغ فى الكلمه أو العبارة أو الجملة أو . . . الخ . ففى الجملة أو الموضوع يحذف جزء أو أكثر وعلى المفحوص أن يأتى به .

قائمة Inventory : أداة تستخدم لسرد كل أو عينة من السلوك أو الاهتمامات أو الاتجاهات أو ... الخ ، تعد مفيدة أو متعلقة بغرض معين . وهى ليست فى العادة مقياس وليست لها إجابات صحيحة أو خاطئة .

قائمة الشخصية Personality Inventory : أداة يقصد منها تقييم أو قياس واحد أو أكثر من الجوانب غير العقلية من التكوين النفسى للفرد .

قدرة عامة General Ability : تعبير عام يستخدم فى حالة القدرات عندما نتعامل مع مشاكلها على نطاق واسع . ويستخدم كمرادف للذكاء . ويرادف فى القياس التربوى القدرة المدرسية .

قدرة خاصة Special Ability : اصطلاح يستخدم مع قدرات من مثل الميكانيكية ، والكتابية ، والموسيقية ، والفنية ... الخ ، متميزة بذلك عن القدرة العامة .

القوة التفريقية Discriminating Power (وقد تسمى القدرة التمييزية) : قدرة فقرة الاختبار على التفريق بين الأشخاص الذين يتمتعون بدرجة مرتفعة من السمة ، من ناحية ، وبين من يتمتعون بدرجة منخفضة من السمة من ناحية أخرى .

القيم Values : تنظيمات معقدة لأحكام عقلية انفعالية نحو الأشخاص أو الأشياء أو المعانى سواء كان التفضيل الناشئ عن هذه التنظيمات صريحاً أو ضمنياً .

قيمة الصعوبة Difficulty value : النسبة المئوية من جماعة معينة ، مثل تلاميذ من سن واحد أو فرقة دراسية واحدة ؛ التى أجابت عن الفقرة إجابة صحيحة . وكلما كانت الفقرة سهلة ، أو كلما زادت النسبة المئوية من التلاميذ الذين يجيبونها إجابة صحيحة ، زادت قيمة صعوبتها .

المئين Percentile : درجة نقطة فى التوزيع تقع تحتها النسبة المئوية من الدرجات التى يشير إليها المئين المعين . فمثلا ، المئين الثلاثين هو النقطة التى تقع تحتها ٣٠٪ من الدرجات ،

المتوسط (م) Mean (M) : وسنقتصر فى التعريف على المتوسط الحسابى فهو أكثر طرق المتوسطات شيوعاً : والمتوسط الحسابى Arithmetic mean هو المجموع الكلى لمجموعة من الدرجات مقسوماً على عدد الدرجات . ويرمز إليه فى كتب الإحصاء الأجنبية بالرمز $\bar{X}$.

المحك Criterion : معيار نحكم به على الاختبار أو نقيمه ، وقد يكون مجموعة من الدرجات أو المقاييس أو التقديرات أو الإنتاج أو ... الخ ؛ صمم الاختبار للتنبؤ بها ، أو للارتباط معها ، كمقياس لصدقها . وهو أيضاً مجموعة من المفاهيم أو الأفكار المستخدمة فى الحكم على محتوى الاختبار عند تقدير مضمونه أو صدقه المنطقى :

المدى Range : المسافة بين أعلى وأقل درجة فى التوزيع .

الاستمرار Continuum : مستقيم ندرك على لمتداده سمة أو مجموعة من الدرجات على أن توزيعها مستمر : فتغير مثل هذا لا يهمنا فيه أن نعرف مدى تقارب قيمتين مثلا ، إذ أن من الممكن دائماً أن تكون قيمة ثالثة بينهما . وقد يسمى المتغير المستمر Continuous variable :

معادلة سبيرمان — براون Spearman-Brown Formula : معادلة تبين العلاقة بين طول الاختبار ومعامل ثباته . وبواسطتها يمكن التنبؤ بمعامل الثبات للاختبار إذا زاد أو نقص طوله بقدر ما . وهى تستعمل كثيراً فى تقدير معامل ثبات الاختبار بأكمله من معامل الثبات بين نصفى الاختبار نفسه .

معادلة كودر — رتشرد سون Kuder-Richardson Formula : وهى معادلة لتقدير ثبات الاختبار من معلومات عن فقرات الاختبار ، أومعرفة الدرجة المتوسطة ، والانحراف المعيارى ، وعدد الفقرات فى الاختبار . وهى تدل على الاتساق الداخلى Internal Consistency للاختبار .

معامل الارتباط Coefficient of correlotion : مقياس لدرجة العلاقة بين مجموعتين من القياسات مأخوذة إما من نفس المجموعة من الأفراد ، أو من ثنائيات من الأفراد .

معامل ارتباط بيرسون Pearson Product-moment correlation : معامل ارتباط يستخدم فى حساب العلاقات المستقيمة على نطاق واسع . وهو يستخدم انحرافات القيم وحواصل ضرب الانحراف على المتغير الأول فى الانحراف على المتغير الثانى بالنسبة لكل فرد .

معامل ارتباط الرتب Rank-Order Correlation : طريقة فى حساب

معامل الارتباط بترتيب درجات كل الأفراد وتحديد العلاقة بينها . ويعتمد على إيجاد الفرق بين رتبتى درجات الأفراد على المتغيرين .

معامل الثبات Reliability Coefficient : معامل ارتباط نحصل عليه بين صورتين متكافئتين من الإختبار أو بين الدرجات فى مرات الإجراء لنفس الاختبار (طريقة إعادة الاختبار Test-Retest) أو بين نصفى الإختبار (الطريقة النصفية Split-half R.) ، أو باستخدام معادلة كودر – ريتشر دسون .

المعايير Norms : قيم تصف أداء مجموعات متعـــددة على اختبار أو قائمة . والمعايير وصفية لأنماط موجودة من الأداء ولا يجب اعتبارها مستويات مثلى أو مستويات مرغوب فى الوصول إليها .

معايير العمر Age Norms : قيم تمثل الأداء الفعلى أو المتوســط لأشخاص من أى مستوى عمر زمنى معين فى أى صفة مقاسة .

التسيع Stanine : درجة معيارية متوسطها ٥ وانحرافها المعيارى تقريباً ٢ وتتراوح بين ١ ، ٩ . وهى أحد الخطوات فى طريقة المقياس ذى النقط التسع ، نتخذها عند تسوية الدرجات المعيارية . وعلى ذلك فالاسم مأخوذ من المعيار التاسع .

معيار الفرقة الدراسية Grade Norm : الدرجة المتوسطة التى يحصل عليها تلاميذ فرقة دراسية معينة .

المعيار القومى National Norm : معيار مبنى على عينة مأخوذة على نطاق الدولة .

المعايير المحلية Local Norms : معايير حســبت بجمع البيانات من مدرسة معينة أو منطقة تعليمية معينة أو فئة حوضية واستخدامها

بدلا من استخدام المعايير القومية أو المعايير الإقليمية ، ذلك كى نقدر أداء الأفراد .

المفتاح المثقب Stencil Key : مفتاح أداء مصمم بحيث أنه عند وضعه فوق ورقة الإجابة Answer Sheet ، وهى إما كتيب الاختبار أو ورقة إجابة منفصلة ، يسمح بالتميز السريع وحصر الإجابات الصحيحة .

مقياس التدرج بالرسم Gradhic Rating Scale : طريقة لتسجيل التقديرات ، تبعاً لقوة الصفة أو السمة على طول خط مستقيم (استمرار) ، وتحت هذا الخط تذكر العبارات الوصفية للسمة المقاسة .

مقياس العمر Age Scale : مقياس تكون وحدات القياس فيه هى الفروق بين مكافئات الأعمار المتتابعة ، ويؤخذ هذا الفرق كما لو كان مكافئاً لأى فرق آخر .

المقاييس المتدرجة Rating Scales : طرق يقوم فيها الشخص المقدر بتسجيل درجة السمة المعينة أو الصفة التى نقيسها . وهى غالباً إما وصفية ، أو رقمية ، أو مقاييس للتدرج بالرسم .

مقياس النقط Point Scale : اختبار يستجيب فيه الفرد عن كل فقرة فيحصل على نقطة أو أكثر إذا كانت إجابته صحيحة . وتتكون درجة الفرد من مجموع هذه النقط .

مكافئ العمر Age Equivalent : هو العمر الزمنى الذى تقدره أو تقدر متوسطة الدرجة المعينة .

مكافئ الفرقة الدراسية Grade Equivalent : مستوى الفرقه الدراسية الذى تقدره أو تقدر متوسطة الدرجة المعينة .

المنوال Mode (وقد يسمى النمط أو الشائع) : هو الدرجة أو القيمة التى لها أكبر تكرار فى التوزيع .

الميول Interests : الميل هو استجابة الفرد استجابة إيجابية أو سلبية نحو شخص أو شىء أو فكرة معينة استجابة لها صبغة وجدانية .

النزعة المركزية Central Tendency : نقطة فى التوزيع تميل أغلب الحالات إلى الوقوع حولها ، والاصطلاح يشير إلى مقاييس هذا الميل . كما يشير إلى القيمة الممثلة للتوزيع . وأهم طرق حسابها هى المتوسط ، والوسيط ، والمنوال .

نسبة التحصيل (AQ) Achievement Quotient : وهى نسبة بين المستوى الفعلى للأداء المدرسى وبين المستوى المتوقع ونحددها بنسبة العمر التعليمى إلى العمر العقلى . أى نسبة التحصيل = $\frac{\text{العمر التعليمى}}{\text{العمر العقلى}} \times ١٠٠$

وتسمى أيضاً بنسبة الإنجاز Accomplishment Quotient . وهذا المفهوم لا يزال مستخدماً وإن كان مضللا ، والسبب فى تضليله أنه يعتمد على افتراض خاطئ مؤداه أن العمر التعليمى والعمر العقلى يرتبطان معاً ارتباطاً تاماً عند التلميذ السوى أو المتوسط .

نسبة التعليم (EQ) Eductational Quotient : وهى النسبة بين العمر التعليمى والعمر الزمنى أى أن :

$$\text{النسبة التعليمية} = \frac{\text{العمر التعليمى}}{\text{العمر الزمنى}} \times ١٠٠ .$$

نسبة الذكاء (I Q) Intelligence Quotient : هى النسبة التى نحصل عليها بقسمة العمر العقلى على العمر الزمنى . أى أن :

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

وهى مقياس للقدرة العقلية العامة يأخذ فى الاعتبار كلا من الدرجة على اختبار الذكاء والعمر الزمنى لمن حصل على هذه الدرجة .

نسبة الذكاء الانحرافية Deviation IQ : درجة معيارية متوسطها ١٠٠ وانحرافها المعيارى حوالى ١٥ تحول إليها نسبة الذكاء .

الوسيط Median (وقد يسمى الأوسط) : هو الشخص الذى حصل على الدرجة التى تقع فى منتصف الدرجات أى أن عدد من حصلوا على درجات أعلى منه يساوى عدد من حصلوا على درجات أقل منه . ودرجته ليست بالضرورة الدرجة المتوسطة .

الفصل الثالث

تصنيف وتعريف الاختبارات النفسية والتربوية وعرض للمشاكل المتعلقة بها .

- مقـــدمة .
- أسس التصنيف .
- التصنيف على أساس ما نقيسه .
- التصنيف تبعاً للمحتوى .
- التصنيف تبعاً لطريقة التصحيح .
- التصنيف على أساس تطبيق النتائج .
- أسس أخرى للتصنيف .
- تعريف بالأنواع الكبرى من الاختبارات النفسية والتربوية .

الفصل الثالث

تصنيف وتعريف الاختبارات النفسية والتربوية وعرض للمشاكل المتعلقة بها .

- مقـــدمة .
- أسس التصنيف .
- التصنيف على أساس ما نقيسه .
- التصنيف تبعاً للمحتوى .
- التصنيف تبعاً لطريقة التصحيح .
- التصنيف على أساس تطبيق النتائج .
- أسس أخرى للتصنيف .
- تعريف بالأنواع الكبرى من الاختبارات النفسية والتربوية .

٢ ــ الذكورة ــ والأنوثة

٣ ــ الترويحى

(و) الاتجاهات والقيم :

١ ــ الاجتماعى

٢ ــ العنصرى Racial

٣ ــ السياسى

٤ ــ الخلقى Moral

٥ ــ العقائدى religious

(ز) اختبارات حسية حركية :

١ ــ الإبصار

٢ ــ السمع

٣ ــ المهارة الحركية

٤ ــ التآزر الحسى الحركى

(ا) البسيط (عين ــ يد)

(ب) المركب (الأطراف والحواس)

٢ ــ تصنيف على أساس شروط الإجراء Administration :

(ا) فردية (الاختبار يعطى لشخص واحد فى المدة الواحدة)

(ب) جمعية

(حـ) التعليمات والفقرات شفوية

(د) التعليمات والفقرات والاستجابات كتابية

(هـ) اختبارات سرعة Speed (تقدم فى زمن حدوده ضيقة لقياس الكمية التى أداها الفرد أداء صحيحاً فى وقت قصير) .

(و) اختبارات قوة (الحدود الزمنية واسعة ولكن درجة الصعوبة هى التى تحدد أقصى ما يصل إليه المفحوص)

٣ ـ التصنيف تبعاً للمحتوى Contént

(ا) الفقرات فى صيغة مقال أو فقرات إجابتها حرة .

(ب) فقرات الاستجابة عليها موضوعية أو محددة :

١ ـ الاختيار من اجابتين :

(ا) صحيح ـ غير صحيح

(ب) نعم ـ لا

(ح) صحيح ـ خطأ

٢ ـ الاختيار من إجابات متعددة

٣ ـ المزاوجة Matching

٤ ـ التكميل ؟

٥ ـ صيغ نادرة .

(ح) اختبارات أدائيـــة أو غير لفظية :

١ ـ عامة

٢ ـ المادة عينة من الشغل فى أعمال معينة .

٤ ـ التصنيف تبعاً للتصحيح :

(ا) اختبارات تتطلب تقييما كيفياً من الخبراء .

(ب) التصحيح بالنسخ . Stencile أى التصحيح بمفتاح مثقب .

(ح) التصحيح بالماكينات

٥ ـ التصنيف على أساس تطبيق النتائج :

(ا) اختبارات غير رسمية (اختبارات يضعها المدرسون) .

(ب) اختبارات مقننة :

٦ ـ أسس أخرى للتصنيف مثل التفريق بين الاختبارات التى :

(ا) تتطلب أقصى أداء يستطيعه الفـــرد (كاختبارات الذكاء والتحصيل و . . . الخ) .

(ب) اختبارات تتطلب أداء عادى أو معتاد (وهى اختبارات الشخصية) .

وكذلك التصنيف حسب التصميم إلى :

(ا) اختبارات ذات بناء محدد .

(ب) اختبارات غير محددة البناء unstructured (كالإسقاطية) والتصنيف على أساس الفلسفة القائم عليها الاختبار .

والتصنيف على أساس نوع المواد فيه مثل :

(ا) لفظية

(ب) أرقام ومسائل حسابية

(حـ) رموز

(د) صور

(هـ) مواقف من الحياة

تصنيف على أساس جوانب السلوك التى تمثلها وتصنيف الاختبارات إلى :

١ ـ اختبارات العامل العام

٢ ـ اختبارات التصنيف العام أو التصفية

٣ ـ اختبارات الاستعدادات الخاصة

٤ ـ اختبارات القدرات الخاصة

٥ ـ اختبارات الاستعدادات الفارقية

وهناك تفرقة أخرى على أساس وسط Médium الاختبار إلى :

١ ـ اختبارات ـ ورقة ـ وقلم Paper-and-Pencil

٢ ــ اختبارات أداء Performance

وهناك تقسيم على أساس المحتوى السائد Content

١ ــ لفظى Verbal

٢ ــ عددى numerical

٣ ــ مكانى Spatial

وهناك تصنيف إلى اختبارات :

١ ــ لغوية language

٢ ــ غير لغوية non - language

وكذا تصنيف على أساس طبيعة الاختبار إلى :

١ ــ إسقاطية projective

٢ ــ مفصلة structured أو ذات بناء محدد .

إلى آخر هذه التصنيفات على أن بكل أساس بعض وجاهته ومبرراته وليس يهم فى الكثير أن يقيم أساس التصنيف إذ أن لا طائل فى هذا . ذلك فضلا عن أن التصنيف الجامع المانع لا وجود له بين الاختبارات النفسية وربما كان هـذا راجعاً أساساً إلى أنها نفسية . فمثلا نجد أن الاختبار قد يستخدم فى غير ما صمم من أجل قياسه . كما أن بعض الاختبارات يمكن استخدامها استخدامات مختلفة من حالة إلى أخرى ومن قائم بالإجراء إلى آخر .

والآن نتحدث بشىء من التفصيل عن هذه الأسس وتلك التصنيفات معرفين بالأنواع الكبرى من الاختبارات النفسية ، وبالأسس التى صنفت تبعاً لها هذه الاختبارات ومتناولين بعض المشاكل المتعلقة بها . على أننا سنعود للحديث المفصل فى فصول قادمة عن أغلب هذه الاختبارات .

وكذلك الحال فى اختبارات اختيار المستخدمين والعمال الذى أصبح بعضها يسمى اختبارات « التصنيف العام » أو « اختبارات التصفية » . والاسم مأخوذ من حقيقة أن اختبارات الذكاء ، وخاصة الجمعية منها ، تستخدم كأدوات للتصنيف أى لتقسيم الجماعة المختبرة إلى مستويات من يجتازها نطبق عليه بعد ذلك اختبارات أكثر تفصيلا للاستعدادات الخاصة ، كالاستعداد الميكانيكى ، والكتابى ، والموسيقى ، والحسابى . وأصبحت الاختبارات متوفرة فى كل هذه الميادين ، تستخدم فى التوجيه التربوى والمهنى ، والاختيار المهنى ، وفى مجالات أخرى كثيرة .

وقد أصبح هناك اعتقاد متزايد لتدعيم ما يسمى باختبارات الذكاء . ولا أدل على هذا من اجتماع ضم ٧٩ من أقطاب علم النفس سنة ١٩٤٤ . وافق ٥٥ منهم على الاعتقاد بأنه « من الممكن إنجاز الكثير لو ركز علماء النفس على عوامل عقلية منفصلة » وليس معنى هذا أن ذلك الفريق لا يعطى أهمية لاختبارات الذكاء ، فعلى النقيض من ذلك وافق ٧٥ ٪ منهم على أنها أصلح الاختبارات فى أغراض تصنيف الناس تبعاً للقدرة العامة كما فى الجيش ، وفى المدارس ، وفى الصناعة .

اختبارات الاستعدادات الخاصة :

ونتيجة لهـذا ظهرت اختبارات الاستعدادات الخاصة Special Aptitude tests . ثم تلاها ظهور اختبارات الاسـتعدادات الفارقية Differential Aptitude tests وإن كان هدف النوعين هو ملء الفجوات التى تركتها اختبارات الذكاء العام . وميادين الاستعدادات الخاصة ميادين متخصصة جداً ، مثل الاستعداد الموسيقى ، والفنى ، والميكانيكى ...الخ ، دعت إليها الحاجة الماسة للاختبار المهنى الذى شجع التنبؤ بهذه الاستعدادات . والاستعداد Aptitude هو إمكانية الوصول إلى درجة من الكفاية أو القدرة

عن طريق التدريب : وفى الاختبارات التى تقيسه يقصد به السرعة التى نتوقع بها أن يتعلم الفرد طرق الأداء المتعلقة بهذا الاستعداد : بمعنى أن الاستعداد للأعمال الكتابية ، مثلا ، يشير إلى قدرة الفرد على تعلم طرق الأداء المكتبية إذا توفر له التدريب .

واختبارات الاستعدادات تتداخل فى وظائفها مع بطاريات الاستعدادات الفارقية ، وإن كان هذا لا يحدث دائماً : وقد توفرت الدراسات المتعددة فى مجال الاستعدادات الخاصة حتى أصبحت هناك مقاييس غاية فى التخصص تقيس جوانب غاية فى البساطة ، مثل حدة الحس أو سرعة حركة الإصبع . وعلى الجانب الآخر نجد اختبارات أخرى معقدة شاملة للاستعدادات مثل الاستعداد للمحاماة أو للهندسة أو للطب :

اختبارات التحصيل بأنواعها :

ويجب أن نشير هنا إلى التمييز الذى قال به البعض بين اختبارات الاستعدادات واختبارات التحصيل . فاختبار التحصيل يقيس أثر نوع معين من التدريب ، ومثله الاختبار التعليمى ، إذ يقنن الاختبار على برنامج دراسة مادة معينة . أما اختبار الاستعداد فيقيس السرعة المتوقعة لتعلم الفرد أو مدى استفادته من التدريب أو التعليم : واختبارات التحصيل فى مجال الصناعة تسمى « اختبارات الكفاية » Proficiency test ، فهى تقيس مدى استفادة الفرد من برنامج التدريب :

أما « اختبارات الحرفة » Trade Tests فهى نوع من اختبارات التحصيل أيضاً فى مجال الصناعة . وتستخدم فى اختيار وتصفية الموظفين والعمال فى الصناعات والحرف . وعلى أى حال فالتفرقة بين الاختبار التحصيلى ، فى المدرسة ، واختبارات الكفاية أو اختبارات الحرف ، فى المصنع ، تفرقة نسبية أكثر منها مطلقة . إذ أن كل نوع من الاختبارات

وكذلك الحال فى اختبارات اختيار المستخدمين والعمال الذى أصبح بعضها يسمى اختبارات « التصنيف العام » أو « اختبارات التصفية » . والاسم مأخوذ من حقيقة أن اختبارات الذكاء ، وخاصة الجمعية منها ، تستخدم كأدوات للتصنيف أى لتقسيم الجماعة المختبرة إلى مستويات من يجتازها نطبق عليه بعد ذلك اختبارات أكثر تفصيلا للاستعدادات الخاصة ، كالاستعداد الميكانيكى ، والكتابى ، والموسيقى ، والحسابى . وأصبحت الاختبارات متوفرة فى كل هذه الميادين ، تستخدم فى التوجيه التربوى والمهنى ، والاختيار المهنى ، وفى مجالات أخرى كثيرة .

وقد أصبح هناك اعتقاد متزايد لتدعيم ما يسمى باختبارات الذكاء . ولا أدل على هذا من اجتماع ضم ٧٩ من أقطاب علم النفس سنة ١٩٤٤ . وافق ٥٥ منهم على الاعتقاد بأنه « من الممكن إنجاز الكثير لو ركز علماء النفس على عوامل عقلية منفصلة » وليس معنى هذا أن ذلك الفريق لا يعطى أهمية لاختبارات الذكاء ، فعلى النقيض من ذلك وافق ٧٥ ٪ منهم على أنها أصلح الاختبارات فى أغراض تصنيف الناس تبعاً للقدرة العامة كما فى الجيش ، وفى المدارس ، وفى الصناعة .

اختبارات الاستعدادات الخاصة :

ونتيجة لهـذا ظهرت اختبارات الاستعدادات الخاصة Special Aptitude tests . ثم تلاها ظهور اختبارات الاستعدادات الفارقية Differential Aptitude tests وإن كان هدف النوعين هو ملء الفجوات التى تركتها اختبارات الذكاء العام . وميادين الاستعدادات الخاصة ميادين متخصصة جداً ، مثل الاستعداد الموسيقى ، والفنى ، والميكانيكى ...الخ ، دعت إليها الحاجة الماسة للاختبار المهنى الذى شجع التنبؤ بهذه الاستعدادات. والاستعداد Aptitude هو إمكانية الوصول إلى درجة من الكفاية أو القدرة

عن طريق التدريب . وفى الاختبارات التى تقيسه يقصد به السرعة التى نتوقع بها أن يتعلم الفرد طرق الأداء المتعلقة بهذا الاستعداد : بمعنى أن الاستعداد للأعمال الكتابية ، مثلا ، يشير إلى قدرة الفرد على تعلم طرق الأداء المكتبية إذا توفر له التدريب .

واختبارات الاستعدادات تتداخل فى وظائفها مع بطاريات الاستعدادات الفارقية ، وإن كان هذا لا يحدث دائماً . وقد توفرت الدراسات المتعددة فى مجال الاستعدادات الخاصة حتى أصبحت هناك مقاييس غاية فى التخصص تقيس جوانب غاية فى البساطة ، مثل حدة الحس أو سرعة حركة الإصبع . وعلى الجانب الآخر نجد اختبارات أخرى معقدة شاملة للاستعدادات مثل الاستعداد للمحاماة أو للهندسة أو للطب .

اختبارات التحصيل بأنواعها :

ويجب أن نشير هنا إلى التمييز الذى قال به البعض بين اختبارات الاستعدادات واختبارات التحصيل . فاختبار التحصيل يقيس أثر نوع معين من التدريب ، ومثله الاختبار التعليمى ، إذ يقنن الاختبار على برنامج دراسة مادة معينة . أما اختبار الاستعداد فيقيس السرعة المتوقعة لتعلم الفرد أو مدى استفادته من التدريب أو التعليم . واختبارات التحصيل فى مجال الصناعة تسمى « اختبارات الكفاية Proficiency test » ، فهى تقيس مدى استفادة الفرد من برنامج التدريب .

أما « اختبارات الحرفة Trade Tests » فهى نوع من اختبارات التحصيل أيضاً فى مجال الصناعة . وتستخدم فى اختيار وتصفية الموظفين والعمال فى الصناعات والحرف . وعلى أى حال فالتفرقة بين الاختبار التحصيلى ، فى المدرسة ، واختبارات الكفاية أو اختبارات الحرف ، فى المصنع ، تفرقة نسبية أكثر منها مطلقة . إذ أن كل نوع من الاختبارات

يمكن أن يستخدم ، فى ظروف معينة ، لقياس أثر الخبرة الماضية والتنبؤ بالإنجاز المقبل .

اختبارات الاستعدادات الفارقية :

وقد أدى التقييم النقدى لاختبارات الذكاء الذى تبع انتشار استخدامها فى كل الميادين إلى توضيح نقطة أخرى ، ألا وهى ، أن الأفراد يختلفون فى أدائهم للأجزاء المختلفة من اختبار الذكاء الواحد . اتضح هذا خصيصاً فى الاختبارات الجمعية حيث تعزل الفقرات عادة فى اختبارات فرعية Subtests . وهذه الاختبارات الفرعية متناسقة المحتوى نسبياً . كما اتضح أن الشخص الواحد يختلف أداؤه جوهرياً من قسم إلى آخر فى الاختبار الواحد ، خاصة إذا كان أحد الأقسام من نوع لفظى وكان الآخر من نوع الرسوم الهندسية مثلا .

وكثيراً ما استخدم الفاحصون ، وخاصة الأكلينيكيون ، تلك المقارنات الداخلية ليصلوا إلى استبصار أكثر بالتكوين النفسى للفرد . وعليه فلم تكن « نسبة الذكاء » أو غيرها من الدرجات الكلية فحسب موضع الدراسة ، بل إن الدرجات على الاختبارات الفرعية أيضاً تفحص فى دراسة حالة الفرد ، ومثل هذا الإجراء لم يعمم ، إذ لم تكن هناك اختبارات ذكاء مصممة بغرض تحليل الاستعداد الفارقى . وإن غلب أن حوت فقرات قليلة على تحديد ثابت لقدرة معينة : وبناء عليه ، لو اختبر الفرد مرة ثانية بصورة أخرى من نفس الاختبار لتغيرت درجاته على الاختبارات الفرعية فيها . فإن أردنا القيام بمثل هـــذه المقارنات فى ذات ــ الفرد ، لاحتجنا إلى اختبارات صممت خصيصاً لتوضـــيح الفروق فى أداء الوظائف المتنوعة .

وقد غطت اختبارات الاستعدادات الفارقية تلك الحاجة ، وهى

نتيجة للتطور فى دراسة تنظيم السمات دفعت إليها دراسات إحصائية عن طبيعة الذكاء والقيام بعمل الارتباطات الداخلية والتحليل العاملى بفضل كيلى T. L. Kelly وثيرستون L. L. Thurstone . وقد أدت دراسات التحليل العاملى Factor Analysis إلى استخلاص عدد كبير من العوامل المستقلة نسبياً أو السمات . وبعضها تمثل بدرجات متفاوتة فى اختبارات الذكاء مثل عامل الفهم اللفظى وعامل الاستدلال الحسابى .

كما أدت تلك الدراسات أيضاً إلى تكوين بطاريات الاستعدادات الفارقية لقياس مستوى الفرد فى كل من هذه السمات . فبدلا من استخراج « نسبة الذكاء » تحسب درجة الفرد فى سمات مثل الفهم اللفظى ، الاستعداد العددى ، الاستدلال المكانى ، الاستدلال الحسابى Arithmatic Reasoning وقد نجحت مثل هذه البطاريات فى تحليل ذات الفرد أى فى التشخيص الفارقى . وهى أيضاً تضيف إلى الميدان معرفة منظمة متعمقة لما لم تتضمنه اختبارات الذكاء فضلا عن اختبارات الاستعدادات الخاصة .

وقد حلت بطاريات الإستعداد الفارقى محل اختبارات التصنيف العام فى عدد من الأغراض مثل توجيه المراهقين والراشدين . وهذه البطاريات تغطى مجالا أوسع من الجوانب النفسية وهى لا تعطى ، كما قلنا ، درجة كلية مثل نسبة الذكاء . وتسمح فضلا عن هذا ، كما قلنا أيضاً ، بعمل مفارقات أو مقارنات بين القدرات الخاصة للفرد الواحد ، ذلك أنها تصل إلى « صفحة نفسية » Psychological Profile . أو « مبيان نفسى » Psychograph . ويعتمد تكوين بطاريات الاستعدادات الفارقية على آخر ما توصل إليه التحليل العاملى ، ما دامت الاختبارات التى تكون هذه البطاريات قد اختيرت لتمثل السمات الأساسية التى حددها التحليل العاملى ، وهى لهذا تضمن تغطية قدرات أكثر تنوعاً دون تداخل .

اختبارات الشخصية :

يرى البعض أن القياس النفسى والتربوى هو قياس للشخصية بكل جوانبها ، وهم فى هذا يستندون إلى فهمهم أو نظرتهم للشخصية باعتبار أنها تتضمن كل الجوانب العقلية ، أى المعرفية ، وغير العقلية التى تتضمن الجوانب المزاجية والانفعالية بل والاجتماعية أيضاً . بينما يرى آخرون أننا نكون أحياناً بصدد جوانب عقلية وأحياناً نكون بصدد جوانب غير عقلية ، وفى الأولى نكون بصدد اختبارات القدرات والاستعدادات والتحصيل ، وفى الثانية نكون بصدد اختبارات تقتصر على المميزات الشخصية مثل التوافق الانفعالى والعلاقات الاجتماعية والدوافع والاهتمامات . . . الخ . ولا يعنى هذا أن الفريق الثانى يغفل العلاقة الدينامية التكاملية بين الجوانب العقلية وغير العقلية ، بل إنهم يرون أن ميزة القياس الأولى هى التحديد ، وهذا فإن الاختبارات التى تقيس الجوانب غير العقليـــة ، وإن كانت هذه الجوانب نفسها تتأثر بالجوانب العقلية وتؤثر فيها أيضاً ، هى ما يسمونه باختبارات الشخصية .

وعلى هذا سنقصر الحديث فى اختبارات الشخصية على ما يقيس منها الجوانب غير العقلية . وقياس الجوانب غير العقلية من السلوك ما زال فى المهد ، إذ أن اختبارات الشخصية هنا تتضمن ، مثلا ، مقاييس للتوافق الانفعالى التى تعرف باختبارات عدم الثبات الانفعالى أو « قوائم الشخصية » Personality Inventories كما تتضمن مقاييس السمات الاجتماعية بما فى ذلك العلاقات الأساســية مع الأشخاص مثل الخضوع ــ السيطرة ، والانطواء ــ الانبساط ، وكذلك الكفاءة ــ الذاتية . وتتضمن أيضاً قياس « الذكاء الاجتماعى » ، وهو المعرفة والمهارات المطلوبة فى كل المواقف الاجتماعية .

وللذكاء الاجتماعى مشكلة تقابلنا عند وضعه تحت أى من التصنيفات . فاختبار الذكاء الاجتماعى ، بما أنه يقيس المعرفة والمهارات التى يقتضيها التوافق فى المواقف الاجتماعية ، هو قياس لمدى القدرة على تكوين العلاقات الاجتماعية من حيث الكم أو العدد ومن حيث العمق أى مدى التأثير المتبادل بين طرفى أو أطراف العلاقة الاجتماعية ، ومن حيث الشدة أى حرارة العلاقة وفتورها ، ومن حيث الدوام أو الثبات أى مدى بقاء هذه العلاقة فى حالة سوية . الذكاء الاجتماعى بهذا المفهوم التحليلى البنائى الوظيفى Functional Structural Analysis (F.S.A.) قد يضم الى اختبارات الذكاء فهو قدرة أو إلى اختبارات الشخصية فهو توافق كما يضم إلى اختبارات الاستعدادات فهو يشير إلى مدى سرعة الفرد فى تكوين العلاقات خلال المواقف .

وتنتمى إلى اختبارات الشخصية أيضاً مقاييس السمات الخلقية مثل الأمانة ، التحفظ ، والتعاونية . وتشمل اختبارات الشخصية أيضاً مقاييس الدفع ، والاهتمامات ، والاتجاهات . والطرق المستخدمة فى اختبارات الشخصية هى الاستبيان ، والطرق الإسقاطية ، وطرق التقدير الذاتى ، والمقاييس المتدرجة ، والاختبارات الموقفية :

وأول سابقة على قياس الشخصية نجدها فى أعمال كريبلين Kraeplein الذى استخدم طريق « التداعى الحر » Free Association مع مفحوصين غير أسوياء ، فكان المثير كلمة يلقيها هو على المفحوص الذى يرد بأول استجابة ترد بخاطره . كذلك اختباراته عن الآثار النفسية للتعب ، والجوع ، والمخدرات . حيث انتهى إلى أن كل هذه الوسائل تزيد التشتت النسبى للتداعى السطحى . كما أن سومر Sommer دعم طريقة التداعى كوسيلة فنية فى القياس ويجب الاعتراف بفضل جالتون ، وبيرسون ، وكاتل ، فى تقنين الاستبيانات Questionnair والوسائل الفنية للتقدير المتدرج إذ أسهمت كل

هذه الوسائل فى دراسة الشخصية ، وإن لم تهدف ، عندما صممت ، إلى قياس الشخصية .

ومن أمثلة استبيان الشخصية ، أو قائمة التقرير الذاتى Self Report Inventory ، صحيفة البيانات الشخصية Personal Data Sheet التى وضعها وودورث Woodworth فى الحرب العالمية الأولى : وكان هدفه تشخيص الحالات العصابية (Neurotic) الخطيرة التى قد تلتحق بالخدمة العسكرية . شملت هذه الصحيفة عدداً من الأسئلة عن الأعراض العصابية العامة يجيب عنها الشخص بنفسه وفى بعضها قوائم فرعية للاضطراب الانفعالى ، ثم قوائم أكثر تخصصاً عن التوافق فى البيت وفى المدرسة وفى العمل : هذا وإن ركزت قوائم تالية ، وضعها آخرون ، على استجابات أكثر تميزاً ، مثل السيطرة — الخضوع . ثم تلتها استبيانات عن الاهتمامات فالاتجاهات .

وهناك نوع آخر من مقاييس الشخصية هو الاختبارات الموقفية ، وهدف هذه الاختبارات لا يعلن للمفحوص ، إذ نلاحظ أرجاعه واستجاباته دون علمه ، قام بأول سلسلة من هذا النوع هارتشون وماى Hartshone and May وقننّاها على تلاميذ المدارس ، فيما يتعلق بالغش ، والكذب ، والسرقة ، والتعاونية ، والمثابرة . كما صممت سلاسل أخرى فى الطيران كان التقدير فيها ذاتياً .

أما الطرق الإسقاطية ، فهى أحد مناهج دراسة الشخصية التى يتعهدها الإكلينيكيون . وبناء مادة الاختبار فيها « غير محدد » Unstructured كى يسمح لذاتية الفرد أن تعمل . والفكرة وراءها هى أن الفرد « سيسقط » طرقه المميزة فى الاستجابة عند التطبيق . وهدف هذه الاختبارات يخفى على المفحوص ، ومن ثم تقل الفرصة أمامه فى أن يخلق متعمداً انطباعاً مرغوباً فيه . وأول أنواعها اختبارات التداعى الحر ، تلتها اختبارات تكوين الجمل :

والطرق الإسقاطية تلجأ إلى الرسم وتركيب اللعب لتكوين منظر ، وإلى الصور الغامضة،وبقع الحبر ، . . الخ : وأكثرها انتشاراً اختبار « الرورشاخ » Rorchash Ink Plots واختبار رسم « المنزل والشـــجرة والشخص » House, Tree, Person (H.T.P.) واختبار « تفهم الموضوع » Thematic Apperception test (T.A.T.)

وأحدث تطور فى إعداد اختبارات الشخصية يعتمد على نتائج التحليل العاملى . وتختار فقرات هذا النوع من بطاريات الاستعدادات ، وتجمع على أساس الارتباطات ونتائج التحليل العاملى . والنمط الغالب منها هو اختبارات الورقة والقلم ، ولكنها لا تزال فى دور الاستطلاع .

وباختصار فإن الأنواع الموجودة من اختبارات الشخصية تعرض مشاكل خطيرة نظرية وعملية أكثر مما تحل ، فحلولها ليست تامة دائماً . ولكل منهج مزاياه الخاصة وعيوبه :

والطرق المختلفة والنظرات المتعددة ونواحى القصور فى اختبارات الشخصية عموماً لا ترجع إلى تقصير الجهود بقدر ما تعزى إلى طبيعة الشخصية ، فإن هذا المفهوم المركب المعقد الفردى الفريد هى التى تخلق أكثر مشاكل قياسها .

والتمييز بين اختبارات القدرة ، واختبارات الشخصية ، برغم ما قلنا ، تمييز مربك مصطنع ، فالفرد عند تطبيق أى اختبار عليه يتأثر بعوامل القدرة وخصائص الانفعال وبدوافعه واهتماماته وبخصائص غير عقلية أخرى . ولكن فى بناء الاختبار عادة تعطى الأهمية إلى نواح فى جوانب السلوك . وتظهر تلك الأهمية فى التعليمات المعطاة للمفحوص وفى تكوين العلاقة الطيبة والتصحيح والتفسير إلى آخر ذلك من عمليات القياس .

ثانياً ــ التصنيف على أساس شروط الإجراء :

ويقصد عادة بشروط الإجراء ما إذا كان الاختبار فنى ، وبالتالى ،

يمكن تطبيقه على المفحوصين أفراداً ، أم أنه قنن وبالتالى ، يطبق على المفحوصين جماعات . ومن شروط الإجراء أيضاً ما يتعلق بالإعداد لموقف الاختبار وبخطوات عملياته وما إذا كان له وقت محدد ، وهنا تصبح السرعة هى العامل الأساسى ، أم أن الوقت فيه متسع كى يجيب من يقدر على الإجابة وهنا تصبح القدرة هى العامل الأساسى : ومن شروط الإجراء أيضاً ما يجب أن يتوفر فى الفاحص وما يجب توفره فى المفحوص وفى الموقف .

وتقسم الاختبارات اعتماداً على هذا الأساس إلى « اختبارات جمعية » Group tests ، و « اختبارات فردية » Individual test . والاختبارات الجمعية تطبق على مجموعة أو مجموعات من الأفراد فى وقت واحد . وهنا يكون موقف القياس جمعياً ويكون المفحوص عضواً فى جماعة . أما الاختبارات الفردية فيطبق الواحد منها على فرد واحد فى المرة الواحدة ، وهنا يكون موقف القياس موقفاً فردياً Individual Setting .

وأبسط فرق ظاهرى بين هذين النوعين من الاختبارات أن التعليمات فى الاختبارات الجمعية أبسط منها فى الفردية . وتعطى التعليمات فى الموقف الجمعى عند بدء الاختبار مرة واحدة فى الجلسة الواحدة . فى حين أن التعليمات ، فى الاختبارات الفردية ، معقدة وكثيرة وهى تختلف من فقرة إلى أخرى فى الاختبار الواحد ، وكل فقرة أو مجموعة فقرات تخصها مجموعة من التعليمات .

وللاختبارات الجمعية ، عادة ، ميزة أخرى هى أن تطبيقها ، من جانب القائم بها أى الفاحص ، لا يحتاج إلى مهارة فائقة وتدريب شاق طويل متخصص ، كما لا يشترط غالباً أن يكون الفاحص إخصائياً نفسياً ، وليس فى هذا خروج على أخلاقيات مهنة القياس : بينما نجد ، فى الوقت نفسه ، أن معظم الاختبارات الفردية يجب أن يتوافر فى القائم بها تدريب شاق طويل متخصص ومهارة فائقة : وفى العادة ، يقوم بها أخصائى نفسى مدرب عليها : وأحياناً ينبغى أن يكون قد درب على الاختبار الذى يطبقه بالذات

فضلا عن إعداده السابق . فإن لم يتوفر هذا كان خروجاً على أخلاقيات المهنة : وإن كان هذا الخروج يحدث أحياناً فى تدريب الإخصائيين النفسيين فى الجامعات ومراكز البحوث .

والاختبارات الفردية تطبق كثيراً فى البحوث الإكلينيكية ، وهنا لا يهتم الفاحص بالكم أى بمدى صحة الاستجابات وبدرجة المفحوص فى الاختبار . بل إنه يهتم أساساً بالكيف أى بطريقة استجابة المفحوص فيلاحظ مدققاً أموراً عدة من مثل زمن الرجع أى الوقت الذى مضى بين السؤال والجواب وبالانفعالات التى صاحبت الإجابة من تعبيرات الوجه وحركة الأطراف ، وبنوع الأخطاء ، فقد يخطئ الفرد فى نوع معين من الأسئلة وينجح فى أخرى تتطلب مهارة أكثر ، وهنا يفترض أن ثمة أسباباً لخطئه . فقد تتعلق الأسئلة التى فشل فيها بخبرات غير سارة أو باضطراب فى القدرة أو غير ذلك من التفسيرات الإكلينيكية .

ولهذا السبب نجد أن الكثير من المؤسسات والهيئات العلمية ومراكز البحوث لا يأخذ بنتائج الاختبار وتفسيره إلا إذا طبقه سيكلوجى درب عليه نظرياً وعملياً ، بل ، وأحياناً اجتاز فيه امتحاناً للتأكد من صلاحيته للقيام بهذا العمل . وفى العادة يجرى الإخصائى النفسى الاختبار فى كراسة منفصلة ويصححه بنفسه ، وفى نفس الوقت ، يقوم أخصائى نفسى خبير أو مشرف عليه بتسجيل إجابة المفحوص فى كراسة أخرى مماثلة ثم يصححها ويفسرها . ولا يؤخذ بتصحيح الإخصائى النفسى المبتدئ إلا إذا اتفق تصحيحه مع تصحيح الإخصائى النفسى الخبير :

وعلى ذلك يجب التفريق بين نتائج كل من الاختبارات الفردية والاختبارات الجمعية . إذ أن الاختبارات الجمعية تطبق فى مجالات تمس الأفراد من حيث هم أعضاء فى مجموعات كبيرة ، فى حين أن الاختبارات الفردية تمس الفرد من حيث هو فرد . ونتيجة لطبيعة وظروف موقف الإجراء ، يرى

البعض أن الاعتماد على نتيجة الاختبار الجمعى أسلم منه على الفردى ، خاصة وأن الاختبار الجمعى فى تقنينه لا يسمح للقائم بإجرائه بالتصرف الشخصى ، فلا تعمل ذاتية الفاحص فى الإجراء ، كما لا تؤثر فى التصحيح الذى يتم بطريقة آلية دائماً ، ودور الفاحص فى تفسير نتائج الاختبار الجمعى محدود إذا قورن بدوره فى تفسير نتيجة الاختبار الفردى .

والمفحوص ، فى الموقف الجمعى ، يكون أقل شعوراً بالموقف منه فى الموقف الفردى . إذ يحس فى الأخير بالحرج ولا يألف الموقف سريعاً فى حين أنه فى الموقف الجمعى يدرك أن غيره يشاركه فيما يمر به ، كما أنه يكون منصرفاً عن الشعور بالموقف لأنه يعلم أن أحداً لا يركز الانتباه عليه . وقيمة هذا أن معايير الاختبار المقنن على مواقف جمعية لا تجعله صالحاً للإجراء فى المواقف الفردية . فإن المعايير المستخلصة بطرق إحصائية تمكننا من مقارنة الأفراد المنتظر قياسهم بنظرائهم من مجموعة التقنين . وعلى هذا نطبق المقياس فى مواقف من نفس نوع المواقف التى قنن فيها . خاصة وأنه فى المواقف الجمعية تبرز عوامل تفتقر إليها المواقف الفردية مثل التعاون ، والتنافس ، والشعور بأن الخبرة عامة يمر بها أفراد مختلفون فى نفس الوقت ، وتوحيد الظروف ، وتوحيد طريقة الإجراء ، . . . الخ .

كما أن هناك عوامل تظهر فى الموقف الفردى يفتقر إليها الموقف الجمعى ، مثل إقامة علاقة طيبة ، وضمان تعاون المفحوص ، وكسب ثقته ، وإثارة اهتمامه ، وإمكان الحصول على استجابات قد لا يمكنه أن يدلى بها فى الموقف الجمعى . هذا وإن كانت هذه المميزات وتلك النقائص تختلف ، من موقف إلى آخر ، ومن اختبار إلى آخر ، ومن فاحص إلى آخر ، كما يرجع الكثير من الفروق إلى نوع الوظيفة أو الوظائف المقاسة وطبيعة المفحوصين . بدليل أن الاختبار الواحد قد يقنن جمعياً مرة وفردياً مرة أخرى : وتختلف المعايير فى الحالتين كما حدث فى اختبار المصفوفات

المتتابعة Progresive Matrices ، إذ اختلفت متوسطات درجات عينتي التقنين وكذلك الانحرافات المعيارية وبالتالى انخفضت الدرجة المتوسطة فى التقنين الفردى عنها فى الجمعى .

وتعـد اختبارات الجيش ألفا والجيش بيتا Army Alphba & Army Beta التى وضعت فى الحرب العالمية الأولى ، تأليف أرثر أوتس Arthur Otis ، من أوائل الاختبارات الجمعية للذكاء . وتطبق الاختبارات الجمعية الآن على نطاق واسع فى الصناعة باسم اختبارات الكفاية ، وفى المدارس تسمى اختبارات التحصيل . أما عن الاختبارات الجمعية للشخصية فإن أشهرها « قائمة مينسوتا المتعددة الأوجه للشخصية » Minnesota Multifasic Personality I nventory (M.M.P.I.) التى وضعها هاثاوى وماككنلى Hathaway and MaKinley بجامعة مينسوتا وتتكون هذه القائمة من عدد من المقاييس التى تقيس أوجها مختلفة للشخصية . وهى تصلح جمعية كما تصلح فردية . وفى الحالة الجمعية يكون الاختبار على شكل كتيب Booklet form أما فى الحالة الفردية فيكون الاختبار على شكل بطاقات Card-form على كل منها أحد أسئلة القائمة وعلى المفحوص أن يسجل فى ورقة الإجابة ما إذا كان مضمون الفقرة ينطبق عليه فيملأ الخانة أمام (نعم) تحت رقم هذا السؤال . وإذا كانت الفقرة لا تنطبق على المفحوص أولا يوافق عليها ، ملأ الخانة (لا) تحت رقم هذا السؤال بالذات : والأسئلة تقيس عشرة جوانب من الشخصية كل جانب يخصه عدد من الأسئلة ، ولكن الأسئلة كلها مرتبة عشوائياً بحيث لا يتبين المفحوص أن مجموعة بذاتها تتناول جانباً بذاته من جوانب شخصيته . وبهذه الطريقة العشوائية أيضا يمكن أن تعاد صياغة السؤال ويتكرر وروده فى مكان آخر ، وذلك للتحقق من مدى اتساق وثبات استجابات المفحوص .

وسنعود للحديث عن اختبار M.M.P.I. هذا تفصيلا فى فصل آخر .

أما جيلفورد وزوجته J.P. Guilford and Ruth E. Guilford فقد قاما بعمليات التحليل العاملى لمقاييس الشخصية ، واستخلصا ثلاثة عشر عاملا . ثم وضعت اختبارات ثلاثة تقيس هذه العوامل ، أولها اختبار S.T.D.C.R. والحروف مأخوذة من اختصارات العوامل الخمس التى يقيسها هـذا الاختبار . وسنعرض لهذه الاختبار والاختبارتين المشتقين من الثمانية عوامل الأخرى (الثانى يقيس خمس عوامل والثالث يقيس ثلاث عوامل) فيما بعد .

وكذلك اختبار الرورشاخ لبقع الحبر الذى تبذل الآن محاولات لتطبيقه جمعياً بعرض بطاقات البقع بواسطة الفانوس السحرى . ومن اختبارات الشخصية اختبار T.A.T. . وأغلب الطرق الإسقاطية ، على العموم ، فردية .

وقضية الاختبارات النفسية الفردية قضية تعرضت للكثير من الادعاءات والاتهامات وإن ثبت بطلان أغلب هذه الدعاوى . وكل هذه الدعاوى تقريباً تتعلق باستخدام الاختبارات النفسـية استخداما فردياً عن يد الإكلينيكيين : وأصحاب هـذه الإدعاءات أساساً هم الأطباء النفسيون وحجتهم فى هذا أن المفحوص فى الموقف الفردى يكون فى حرج وأنه يكون واعياً بذاته بدرجة تعطل استجاباته الطبيعية ، أى التى تمثل تكوينه النفسى . والحق أنهم فى إنكارهم هذا يستندون إلى إطار مرجعى معين فى التفكير يفضى بهم فى نهاية الأمر إلى إنكار قيمة القياس كله فهم ضعيفو الإيمان به : ودفع هذا الادعاء يأتى من أننا هنا نقارن المفحوص بمفحوصين كانوا ، فى جماعة التقنين ، يمرون بنفس الموقف تقريباً ويشعرون بالمعاناة من ذات مشاعر الوعى الشديد بالموقف وبالتحرج إلى آخر هذه المشاعر .

فهنا يصدق على الفرد ما صدق على من وضع فى موقف مشابه : وأساس المقارنة إذاً هو نفس أساس المقارنة فى كل عمليات القياس ، ألا وهو توحيد الظروف . ومقياس المقارنة هو هو انحراف الشخص عن معيار معين انحرافاً ذا دلالة على أن موضعه النسبي بعيد عن متوسط موضع من مروا بالموقف نفسه ، وأن هذا البعد يقارن بمتوسط بعد أفراد عينة التقنين عن متوسط درجات أفرادها (ومقياس المقارنة إذا هو المتوسط ، م ، والانحراف المعيارى ، ع) .

ومن الأمثلة الواضحة على الاختبارات النفسية والتربوية من النوع الفردى فيما يتعلق بالذكاء والقدرة العقلية مقياس « ستنفرد — بينيه » مراجعة سنة ١٩٣٧ التى أعدها تيرمان ، ميريل . وكذلك مقياس « وكسلر بلفيو لذكاء الراشدين والمراهقين » الذى وضعه وكسلر وهما أفضل وأشهر المقاييس العالمية لقياس القدرة العامة . أما عن اختبار « المصفوفات المتتابعة » الذى وضعه راڤن فهو فردى ، إذ قنن فى مواقف فردية ، وجمعى إذ قنن فى مواقف جمعية ؛ وله صـــورتان أولاهما للأطفال والأخرى للراشدين ؛

ومن الاختبارات التى استطاعت أن تفرق بين القدرات المعرفية ، اختبار « فيرنو » الذى وضعه الدكتور فيرنو Furneaux ، وهو اختبار جمعى ميزته التفريق بين القدرات المعرفية فى ناحيتين هما : الأولى هى الدقة فى حل المقياس ، والثانية هى السرعة فى العمليات العقلية . وللعاملين أهميتهما فى التقييم والتشخيص والاختيار والتوجيه .

ونعود إلى الحديث عن الاختبارات الفردية فى مقابل الجمعية ، فالاختبارات الفردية مثل « ستنفرد — بينيه » تتطلب مهارة عالية من الفاحص وهى مصممة إكلينيكيا أساسا للبحث المتعمق للحالات الفردية ؛ وهى تمكن

الفاحص من الملاحظة الدقيقة . إذ يلاحظ ، وقد يسجل ، نوع الأداء ، والاستجابات الاجتماعية والانفعاليـــة للمفحوص . والفاحص هكذا يعرف أكثر من مجرد « نسبة الذكاء » عن المفحوص ، خاصة إذا كان اخصائياً نفسياً خبيراً . والاختبار الفردى زيادة على ذلك يمثل مقابلة مقننة فيوفر فرصة لإقامة علاقة طيبة وللتعاون والاحتفاظ باهتمام المفحوص بالاختبار والموقف . كما أن الفاحص يتحاشى أى ظروف معطلة للمفحوص .

وكثير من فقرات الاختبارات الفردية يحتاج من الفحوص إلى استجابة شفوية أو معالجة يدوية . ومعظمها يلزمه التوقيت الفردى عند الإجراء . كما أن هذا النوع من المقاييس نشأ عن اهتمام بحالات الاضطراب فى القدرة العقلية والتخلف المدرسى ، ولهذا فهى أدوات إكلينيكية . بينما نشأت وتطورت الاختبارات الجمعية بفضل الحاجة العملية المتزايدة إلى اختبارات للتصنيف والتصفية وحصر الإمكانيات العقلية وتخطيط القوى البشرية . وقد حدث مثلا سنة ١٩١٧ أن طلب من علماء النفس الإسهام بجهودهم فى الحرب العالمية الأولى فأعد آرثر آوتس Arthur S. Otis اختبارين جمعيين هما Army Alpha و Army Beta . الأول يستخدم فى عمليات القياس العادية مع مفحوصين متعلمين يعرفون الإنجليزية ، أما الثانى فيطبق على الأميين والأجانب الذين لا يعرفون الإنجليزية . والاختباران جمعيان يمكن تطبيقهما على جماعات كبيرة . عدل الأول بعد الحرب العالمية الأولى كى يناسب الشئون المدنية .

والاختبارات الجمعية مبسطة التعليمات وخطوات الإجراء ، بحيث لا يصعب الإجراء على الفاحص المتوسط التدريب . وإن كان انتشار استخدام الاختبارات الجمعية قد أدى إلى تطوير حركة القياس النفسى والتربوى وخاصة فى مجال الجيش والمدرسة ، إلا أن سهولتها النسبية فى الإجراء والتصحيح وشيوع استخدامها قد جلبا على حركة القياس التورط

والتعرض لكثير من النقد ، إذ أصبحت الاختبارات (موضة) ، وأصبح فى وسع من يريد قياس ابنه أو عائلته ؛ حتى لو لم يكن ذا علم بأى الأسس النفسية ، أصبح فى إمكانه أن يقوم بهذا ، فالأمر بسيط . كما أصبح فى إمكان بعض ذوى الأغراض الشخصية أن يجروا الاختبارات الجمعية فيثبت أحدهم مثلا أن متوسط ذكاء الجزائريين منخفض عن متوسط ذكاء الفرنسيين ، كما قام إنجليزى آخر بإعلان أن فئة من شعوب أواسط أفريقيا تقل فى ذكائها وفى مستوى عملياتها العقلية كثيراً عن أبناء أوروبا البيض ! !

أدى استخدام الاختبارات الجمعية على نطاق واسع إذاً إلى الوصول إلى ما رفضه العلم وأثبتت خطأه البحوث التجريبية لعلماء من نفس الجنسيات (الفرنسية والإنجليزية) . كما أدى هذا إلى أن يدعى بعض المشتغلين بحركة القياس ، وخاصة غير المتعمق منهم إلى أن يؤمل فى الكثير من الاختبارات وإلى أن يعلن للعامة أن مختلف جوانب الحياة العقلية والنفسية يمكن تقديره بدقة فائقة : وهكذا ، توقع العامة من الاختبارات أن تحقق أكثر مما تستطيع ففشلت فى مسايرة توقعاتهم ، بل وأصبحوا يعممون هذا الفشل على كل الاختبارات النفسية : فهى فى نظرهم لا تستطيع شيئاً وتتكلم عن مفهومات غامضة لا تستطيع تقديرها تقديراً ثابتاً دقيقاً .

اختبارات القوة واختبارات السرعة

يفرق بين الاختبارات النفسية والتربوية على أساس شروط الإجراء أيضاً تبعاً للعامل الأساسى الذى تتوقف عليه طبيعة استجابة المفحوص . فالاستجابة مرة تعتمد على عامل السرعة speed ومرة أخرى تعتمد على عامل القوة power . ففى اختبارات السرعة يفرق بين الأفراد تبعاً لسرعتهم فى الأداء ، ولهذا فاختبار السرعة يتكون من فقرات قليلة الصعوبة ، وقد يتكون من فقرات متساوية فى الصعوبة تقريباً بحيث يستطيع المفحوص أن

يحلها جميعها . ولكن عدد هذه الفقرات أكبر من أن يحلها المفحوص فعلا فى حدود الزمن الذى وضع للاختبار. والزمن هنا محدد بحيث لا يكون فى إمكان أحد أن يجيب عن الفقرات كلها . ومن هنا تكون درجة الفرد وهى عدد الأسئلة التى أجاب عليها إجابة صحيحة ممثلة لسرعته فى أداء ما يستطيعه ؛ وعلى هذا تكون اختبارات السرعة هى اختبارات فى الدقة والسرعة معاً فى أداء مستوى خاص من الأسئلة .

وعلى الطرف الآخر نجد اختبارات القوة Power tests وهى اختبارات أسئلتها متدرجة فى الصعوبة والغرض منها قياس أعلى مسـتوى يستطيع المفحوص أن يجيب عليه وتكون درجة المفحوص هى عدد الأسئلة التى أجاب عليها إجابة صحيحة . ويراعى فى اختبارات القوة أن يكون الزمن كافياً لكل مفحوص أن يصل إلى أعلى مستوى يستطيع الإجابة عليه إجابة صحيحة . لذلك فاختبارات القوة تقيس قوة المفحوص فى المادة موضوع الاختبار ولكنها فى نفس الوقت تقيس دقة المفحوص فى الإجابة عن هذه الأسئلة.

أى أن الدقة تدخل على أى حال فى كل من اختبارات السرعة واختبارات القوة . وبعبارة أخرى تكون الدرجة التى يحصل عليها المفحوص فى اختبار القوة دليلا على أقصى مستوى صعوبة يستطيع الإجابة عليه إجابة صحيحة .

والمثال التالى يوضح الفرق بين اختبارات السرعة واختبارات القوة ؛ فاختبار السرعة يقابل سباقاً وضعت فيه موانع كثيرة جميعها من ارتفاع صغير يستطيع جميع المتسابقين أن يتخطوها ، وتقاس هنا سرعة المتسابقين بعدد الموانع التى استطاعوا أن يتخطوها فى زمن محدد ؛

أما اختبار القوة فهو كالسباق الذى وضعت فيه موانع متدرجة فى الارتفاع وتصل إلى ارتفاعات لا يستطيع أى واحد من المتسابقين أن يجتازها . ففى هذا السباق تقاس قوة المتسابق بأعلى ارتفاع من الموانع

استطاع اجتيازه بنجاح ويكون ارتفاع هـــذا الحاجز هو أقصى قدرته على القفز .

ويجب أن نلاحظ هنا أن غالبية الاختبارات الشائعة فى التحصيل وفى الذكاء وفى القدرات تجمع بين خصائص اختبارات السرعة واختبارات القوة فهى تتكون من أسئلة متدرجة فى الصعوبة ولكن الدرجة تمثل مقدار ما أجابه المفحوص إجابة صحيحة بصرف النظر عن صعوبة الأسئلة المجابة .

ونحن نجد فى اختبارات القوة أن الزمن المحدد يسمح لأى مفحوص أن يحاول الإجابة عن كل الفقرات إن كان يقدر أى إن كان مستواه فى الوظائف التى يقيسها الاختبار مرتفعاً بحيث ينهى الإجراء وتكون إجاباته صحيحة . والفقرات هنا متدرجة الصعوبة . ولكن غالباً ما يستحيل ، نسبياً ، حل بعضها بحيث لا يحصل المفحوص على الدرجة النهائية . فالمفحوص إذاً فى اختبار السرعة لا يحصل على الدرجة النهائية وإن كان مستواه فى الوظائف التى يقيسها الاختبار مرتفعاً بدرجة تمكنه من أن يحل كل الفقرات ، ولكنه لا يحلها جميعها فعلا لأن الزمن المحدد يحول دون ذلك . والمفحوص فى اختبارات القوة ، أيضاً ، لا يحصل على الدرجة النهائية ، ولكن ليست السرعة فى حل الفقرات فى حدود الزمن هى التى تحول دون ذلك كما فى اختبارات السرعة ، ولكنه مستوى قدرته الذى يحول دون ذلك .

والفكرة فى ذلك هى أن واضع اختبار السرعة لا يستطيع معرفة أقصى سرعة لمن سيطبق عليهم الاختبار ، ولهذا فهو يضع عدداً كبيراً من الأسئلة يستبعد أن يجيب عنه أى مفحوص إطلاقاً فى حدود الزمن المعين للاختبار . وهنا يكون سقف الاختبار كافياً لتمكين كل فرد من أن يظهر كل ما يستطيع إنجازه من الفقرات من حيث العدد . أما واضع اختبار القوة فهو لا يعرف بالتحديد مستوى قدرات من سيجرى عليهم الاختبار ، ولهذا فهو يضع أعلى

مستوى من الصعوبة . ويكون سقف الاختبار هنا كافياً لتمكين كل فرد من أن يظهر كل ما يستطيع إنجازه من الفقرات من حيث الصعوبة .

هـــذه التفرقة بين اختبار السرعة واختبار القوة تفترض أن بعض الاختبارات خالص ونقى من القوة والبعض الآخر خالص ونقى من السرعة ؛ فهى تفرقة كمية إذ أن أغلب الاختبارات يجمع بين الأساسين بدرجات متفاوتة . وتلزمنا معرفة السرعة والقوة ومدى تأثيرها فى أداء المفحوص للاختبار المعين عندما نفسر درجاته . كما يلزمنا هذا عند التقييم الفنى للاختبار ، إذ أن تجاهل عامل السرعة قد يوقع فى خطأ كبير عند حساب ثبات الاختبار reliability كما سنعلم فيما بعد .

ثالثاً ــ التصنيف تبعاً للمحتوى :

هناك اختبارات تكون فقراتها فى صيغة مقال essaay وتكون الإجابة عنها حرة free كبعض اختبارات القراءة التى نطلب فيها من المفحوص أن يقرأ مقالا ويلخص ما فهمه . وكذلك الاختبارات التى نطلب فيها من المفحوص أن يقرر أهم خصائص موقف نعرضه عليه كتابة . أو تلك التى يطلب فيها حل مشكلة . وقبل أن نستطرد فى هذا الحديث ، نعرف المحتوى فى هذا المجال بالذات ، بما يطلب من المفحوص أى بطريقة الاستجابة : وطريقة الاستجابة قد تكون ذاتية كما فى اختبارات المقال ، أو موضوعية كما سنشير حالا ، أو أدائية أو غير لفظية .

واختبارات حل المشاكل اختبارات تطلب الإجابة الذاتية فى صورة حرة عن فقرات موضوعة فى صيغة مقال . وهى تعقد للمرشحين لشغل الوظائف التى تتطلب تصرفاً شخصياً ، وهى عادة وظائف عليا من وظائف الإدارة ، كذلك عند اختيار أفراد يقومون بالمقابلة أو يتعاملون مع الجمهور ، أو إخصائين فى شئون الأفراد . ويتمثل هذا النوع أيضاً فى بعض اختبارات الإنشاء إذ يطالب المفحوص بالتلخيص والتركيز وأحياناً بعرض الرأى .

ونجد أمثلة لاختبارات المقال فى بعض الاختبارات المهنية التى تستخدم فى ميدان الإرشاد النفسى والتربوى والمهنى ، حيث يقرأ المفحوص بعض الفقرات ويكتب عن تأثره بها أو عن تأثره بنوع مماثل من خبراته الماضية :

والنوع الآخر من الاختبارات النفسية المقابل لاختبارات المقال هو الاختبارات التى تكون فقراتها موضوعية وتكون الاستجابة عنها موضوعية ومحددة . كأن يجيب المفحوص بصحيح أو خطأ ، أو موافق أو غير موافق ، أو مؤيد جداً — مؤيد — محايد — معارض — معارض جداً كما فى اختبارات الاتجاهات ، وقد يجيب المفحوص بصحيح أو خطأ فى مقاييس الحكم . ومن قبيل التحديد فى الاستجابة نجد اختبارات يجيب المفحوص فيها باختيار إجابة واحدة من بين عدة إجابات موضوعة أمام السؤال . وقد تكون الإجابة بالمزاوجة matching فتوضع مسائل متتالية ، مثلا ، فى صف وتوضع حلول هذه المسائل فى صف آخر . والحلول ليست مرتبة بحسب ترتيب المسائل وعلى المفحوص أن يختار من بين الإجابات إجابة معينة يعتقد أنها صحيحة لتناسب أسئلة معينة وهكذا .

هذه الطرق من الاستجابة ، موضوعية كانت أم ذاتية ، كلها لفظية . وهناك استجابات أدائية . والاختبارات النفسية من نوع الأداء قد تكون عامة مثل اختبار المتاهات الذى وضعه پورتيوس Portius . وقد تكون خاصة فتكون مادة الاختبار الأدائى عينة من الشغل فى أعمال معينة كاختبارات السائقين واختبارات عمال النسيج .

وهناك صيغ تخرج عن هذا ، مثل اختبارات اكتشاف السخافات . وفيها يطلب من المفحوص أن يقرر ما إذا كان هناك وجه للغرابة فى أحد الصور أو القصص كما يسأل أن يحدد عن الشيء الذى يستبعد أن يكون معقولا فى الصورة أو القصة . ومن أمثلة الاختبارات التى تخرج عن هذه الأنواع ما يتبع فى الطرق السيكوفيزيقية Psychophysique إذ نعرض على المفحوص

بطاقات على كل منها خط مستقيم ويطلب منه أن يقدر أطوالها وأن يقارن بينها .

رابعاً ـ التصنيف تبعاً لطريقة التصحيح :

بعض الاختبارات النفسية والتربوية يتضمن فقرات فى كراسة التعليمات يعرف الفاحص الإجابة الصحيحة عنها دون مشقة من مفتاح التصحيح scoring key وأسهلها ، بالنسبة للفاحص اختبارات المسائل الحسابية . ويتطلب البعض الآخر من الاختبارات أن يقوم الفاحص ، بعمليات حسابية لكى يفسر درجة الفرد الخام بناءاً على ما تمده به قائمة المعايير . كما أن هناك اختبارات تسجل الإجابة عنها فى ورقة منفصلة معدة لذلك تسمى استمارة الإجابة Answer sheet . وفى هذه الورقة مسافة تحت رقم كل سؤال يملؤها المفحوص ، إما بكتابة الجواب وإما بوضع علامة × على الحرف الدال على الجواب الصحيح من الإجابات المعطاة مع السؤال ، وقد يقوم المفحوص بملء أحد فراغين أمام أو تحت رقم السؤال أحدهما صحيح والآخر خطأ أو أحدهما نعم والآخر لا ، أو أحدهما موافق والآخر غير موافق . ومفتاح التصحيح هو الأداة التى يكشف بها الفاحص عن الإجابات الصحيحة أو نوع الإجابة . ويصمم هذا المفتاح بحيث تكون به ثقوب إذا طبق المفتاح على ورقة الإجابة تطابق كل ثقب مع الإجابة الصحيحة عن كل سؤال ، فإذا لم يكن تحت الثقب فى ورقة الإجابة علامة تدل إجابة المفحوص ، كانت الإجابة عن السؤال خاطئة أو غير مطلوبة . وبجمع عدد العلامات التى تظهر خلال الثقوب نحصل على الدرجة الخام للمفحوص . وقد يكون للاختبار الواحد عدة مفاتيح ، كل مفتاح منها يخص عدداً من الأسئلة وضع لقياس أحد الجوانب . فإذا فرض أن الاختبار يقيس خمسة جوانب من الشخصية ، صممت له إذاً خمسة مفاتيح للتصحيح مفتاح لكل جانب . فالأسئلة رقم ١ ، ٥ ، ٩ ، ١٠ ، ١٨ . . الخ مثلا تقيس الانطواء .

ومفتاح الانطواء يشمل ثقوباً تطابق مواضع إجابات هذه الأسئلة حسب أرقامها وتحجب ما عداها من الأسئلة . ومجموع عدد الإجابات التى تظهر من الثقوب هو الدرجة الدالة على الانطواء لهذا الفرد .

وهناك تطوير أكثر حداثة فى تصحيح الاستجابات وفيه نلجأ إلى طرق آلية أو ماكينات . وتستخدم هذه الطريقة فى الاختبارات الجمعية التى تطبق على نطاق واسع كما فى اختبارات اختيار الجنود وتصنيفهم . وعلى النقيض من هذا تماماً نجد اختبارات يتطلب تصحيحها تقييما كيفياً لنوع الأداء ودلالته ومصاحباته . وهذا التقييم يقوم به خبير فى اختبار معين بذاته مثل اختبارى الرورشاخ واختبار H.T.P. . وكل الاختبارات الإسقاطية من هذا النوع إذ يعتمد التصحيح والتفسير ، إلى حد كبير جداً ، على خبرة الإخصائى النفسى وتكوينه العلمى وذاتيته . وهى لهذا اختبارات أقل موضوعية مما سبق .

خامساً ــ التصنيف على أساس تطبيق النتائج :

وتبعاً لهذا الأساس من التصنيف تقسم الاختبارات النفسية إلى ، غير رسمية informal Tests واختبارات مقننة standardised T. . والأولى اختبارات توضع عن برامج معينة كالتى يضعها المدرسون عن البرامج الدراسية والتى يضعها الإخصائيون المهنيون عن برامج تدريب مهنى معين . وهى اختبارات غير مقننة . أما الاختبارات المقننة ، وهى التى تهمنا غالباً كإخصائيين نفسيين أو تربويين ، فقد جربت على مجموعات من الأفراد يمثلون نوع ومستوى ونسب الأفراد الذين ستجرى عليهم فيما بعد . وتفسر درجة المفحوص على أساس درجات العينة المستخدمة فى التجريب أى عينة التقنين . وسيأتى المجال الذى نفصل فيه هذا الحديث .

سادسا ـ أسس أخرى للتصنيف :

لا نقصد هنا أن نحصر كل التصنيفات ولا حتى أن نسرد أسسها المختلفة ، ولكن لا بأس من أن نتحدث عن بعض تلك التصنيفات ولو بطريقة موجزة :

يفرق كرونباخ بين اختبارات القدرة Ability (وهى عنده اختبارات أقصى أداء ممكن Maximum Performance) وبين اختبارات الأداء المعتاد (وهى عنده اختبارات الخصائص العادية Habitual Characteristics) . والأولى من نوع اختبارات القدرة كاختبارات الذكاء والاستعدادات الخاصة . أما اختبارات الأداء المعتاد فهى اختبارات الشخصية . واختبارات أقصى أداء تهدف إلى معرفة مستوى أداء الفرد فى أفضل ظروف ممكنة إذ يبذل كل جهد ممكن فى مثل هذه الاختبارات لضبط الدفع والاهتمام والظروف المحيطة والعوامل الأخرى بحيث تضمن أن الفرد سيبذل أقصى جهد فى الاختبار . بينما نتوقع من الفرد ، فى اختبارات الشخصية ، استجاباته المعتادة أو العادية . فاختبارات الشخصية لا تبحث عما يعتقد الفرد أنه أحسن حل ، أو الحل الذى يود أن يقوم به ، بل تبحث عن سلوكه العادى فى الموقف المعين .

وكذلك التصنيف إلى اختبارات ذات بناء محدد structured tests واختبارات بناؤها غير محدد Unstructured T. وتتراوح درجة البناء فى الاختبارات النفسية بين مجـــرد تحديد معالم صوره تعرض على المفحوص ، من ناحية ، وبين موقف مصمم بدقة وعناية فائقة فيه عوامل مركبة ؛ ومواد ، وطرق ، وعقبات ، وأشخاص بل وانفعالات كما فى الاختبارات الموقفية . والاختبارات الإسقاطية اختبارات بناؤها غير محدد ، واختبار T.A.T. عشرون بطاقة على كل منها صورة

غير واضحة وضوحا تاما . وهذه الدرجة من البناء أعلى مما نجدها فى اختبار الرورشاخ إذ أن المادة فيه بقع حبر لم يقصد أنها تمثل شكلا ما ، وليس عدم تحديد بناء الاختبار بدقة عيبا فى الاختبارات الإسقاطية بل إنه نوع من التصميم قصد منه إفساح الفرصة لذاتية الفرد وإسقاطاته وتنظيمه للموقف أو إدراكه له وفى هذا ما فيه من نواحى الإسقاط والتعبير الحر إلى آخر الجوانب اللاشعورية .

والفرق الأساسى بين الاختبارات الإسقاطية والاختبارات المفصلة هو من حيث طبيعة المادة التى يحتويها الاختبار كمنبه يثير عمليات ووظائف نفسية معينة من ناحية ، والمنبه فى الاسقاط إجمالى global من ناحية أخرى . فالمفروض فيها أنها تؤثر من حيث الكل لا من حيث تفاصيلها . وهناك فرق آخر بين النوعين ، هو أن الاخصائى النفسى عندما يتناول نتائج تطبيق الاختبار الإسقاطى ينظر إلى الفرد ككل ، فى حين أنه فى الاختبارات المفصلة يؤكد على أجزاء من هذا الكل .

كما أن الاختبارات النفسية والتربوية تصنف على أساس الفلسفة التى قام عليها كل منها ، وكذلك على أساس نوع المواد التى تستخدمها : فهناك فقرات لفظية ، وفقرات من أرقام ومسائل حسابية ، أو رموز ؛ وأخيرا هناك اختبارات تأخذ موادها من مواقف الحياة :

وهناك تفرقة بين الاختبارات النفسية على أساس وسط الاختبار medium ؛ إلى اختبارات ورقة وقلم واختبارات أداء : بدأ هذا التقسيم فى اختبارات الذكاء ثم انتقل إلى اختبارات الشخصية ، وفى اختبارات الورقة والقلم يعطى لكل مفحوص نسخة من الاختبار يجيب عنها بنفسه ، وتترك للإجابة مسافات فارغة فى كراسة الأسئلة أو فى ورقة إجابة منفصلة : والمفحوص فى الاختبارات الأدائيــة يقوم بمعالجة الأشياء

يدويا ، أو بمعالجة الصور ، أو المكعبات ، أو الأجهزة الميكانيكية ، أو قد يقوم بنشاط معقد يمثل بعض أوجه النشاط فى مواقف الحيـــاة اليومية . والاختبارات الأدائية فردية غالبا ، إذ يصعب توفير مجموعة من الأجهزة والمواد تكفى للإجراء الجمعى ، كما يصعب ضبط الموقف ويتكلف الكثير وتعطل القدرة ويحتمل الغش بالمحاكاة . هذا وإن كانت اختبارات الأداء تطبق أحيانا على مجموعات صغيرة من الأشخاص ، وهنا توضع حواجز على مائدة الاختبار كى لا يغش المفحوصون . وهذا الاحتياط لا لزوم له فى حالة الاختبارات الحسية — الحركية المعقدة إذ يصعب الغش وكذا لا يتأثر المفحوصون ببعضهم البعض فى الأداء .

وهنـــاك تصنيف إلى اختبارات لغوية Language وغير لغوية Non Langunge . وفى الأخيرة لا تستخدم أى لغة كتابة أو كلاما لا فى التعليمات أو الفقرات . وهى تستخدم فى حالة الأميين والأجانب والصم والعجزة كما تستخدم فى حالة الاختبارات التى تقيس السمات التى يفترض أنها موحدة بين مختلف الثقافات ويسمى هذا النوع من الاختبارات « الاختبارات عبر الحضارية .Cross Cultural T » . وقد تكون الاختبارات غير اللغوية من نوع الورقة والقلم بالصور ، والرسوم ، والرموز ، ورسوم المكعبات . وهنا يسجل المفحوص الإجابة برموز بسيطة والتعليمات تعطى بالإشارة والحركات وعرض الخرائط والرسوم . ولكن الاختبارات غير اللغوية فى أغلب الأمر لا تستغنى عن اللغة ، فاللغة إن لم تستخدم فى التعليمات أو فى الفقرات استخدمت فى الإجابة — فهى على أى حال مستخدمة وإن لم تكن عنصرا يحـــدد الإجابة أو يقف دون تعبير الفرد عن استجابات تمثل قدرته .

وهناك تقسيم بين الاختبارات على أساس المحتوى الذى يسود الاختبار المعين . وهذا المحتوى السائد قد يكون لفظى ، أو عددى ، أو مكانى و . . الخ . وفيها قد تستخدم الصور لقياس الفهم اللفظى كما فى اختبارات ما قبل المدرسة Preschool إذ ينطق الفاحص بكلمة أو عبارة أو جملة تعبر عن شىء صور فى الاختبار ويعلم المفحوص على الصورة التى تناسبه . وهذه الجمل التى يقولها الفاحص قد تصف الشىء أو تسرد أهميته أو تعرف وظائفه وعلى ذلك فاستجابة الفرد باختيار الصورة فى كل هـــذه الحالات تختلف بحسب نوع الفقرات والتعليمات . وهناك اعتراض على هذا النوع من الاختبارات هو أن المفحوصين يجرون نفس العمل بطرق مختلفة ، كما أوضحنا فى حالة الصور ، تعكس درجات مختلفة من نمو الوظائف النفسية المختلفة . ومع هذا ، ما دامت الإجابة واحدة ، يحصل المفحوصون على درجة واحدة مع أنهم اعتمدوا على عمليات عقلية بعضها معقد والآخر بسيط ، والنتيجة على الاختبار واحدة رغم هذه الاختلافات .

قائمة مراجع القسم الأول

1— Adkins, Dorothy C. The effects of Parctice on intelligence test Scores. J. educ. Psy. 1937.

2— American Psychological association. Technical recommendations for Psychological tests and diagnostic techniques Washing. ton : Amer. Psychol - Assoc. 1954.

3— American Psychological Association. Ethical standards of Psychologists. Washington ; Amer. Psychol. Assoc. 1953.

—4 Anastasia, Anne. Psychological Testing N. Y. : Macmillan Co., 1959.

5— Anastasia, Anne. The Concept of Validity in the interpretation of test scores. Educ. Psych : Measurement, 1950.

6— Anastasia, Anne. Differential Psychology (3rd Ed.) N.Y. : Macmillan, 1958.

7— Baron, D. and Bernard, H.W. Evaluation Techniques for Classroom Teachers. Mc-Grow Hill Book Co. 1958.

8— Bean, K. Construction of Educational and Personnel. Mc. Graw - Hill Books Co. 1953.

9— Boring. E. G. A. History of Experimental Psychology. (Rev. Ed.) N.Y., Appleton—Century—Crops, 1950.

10— Cronbach, L. J. Eessntials of Psycological - Testing. N.Y. : Harper 1949.

11— Crow, L. E. and Crow, A. Educational Psychology. American Book Co. (Revissd ed.) 1958

12— Dixon, W. J. and Massey, F. J. Mc. Grow - Hill 1951.

13— Garrette, H. E. Statistics in Psychology and Education. (Third edition) Longmans, Green and Co. 1948.

14— Good, C. V. and Merkel, W. R. Dictionary of Education. (2nd edition) Mc. Graw - Hill 1959.

15 – Goodenough Florencel. Mental testing. N.Y. Rinehart, 1949.

16 – Greene, H.A. and Jorgenson, A.N. and Jerberich, J. R. Longmans, Green and Co. 1947.

17 – Greene, E.B. Meosurement of Human Behavior N.Y. Odyssey 1941.

18 – Guilford, J.P. Fundamental Statistis in Psychology and Education. Mc Grow–Hill Book Co. 1942.

19 – Hilderth, Gertrude H.A. Bibliography of Mental tests and rating (2 nd Ed) N.Y : Psychology Corp. 1939.

20 – Hilderth, Gertrude H.A. Bibliography of mental tests and rating Scales, 1945 supplement N.Y. Psychology. Corp. 1946.

21 – Johnson, V.O. Statistical Methods in Research. Prentice-Hall., Inc. 1949.

22 – Mc. Nemar, Q. Psyclogical statistics. John Willy, Son. 1949.

23 – Mursell, J. L. Psychological Testing. (2 nd edition) N.Y. Longmans, Green and Co. 1950.

24 – Murphy, G. An historical introduction to modern Psych. Rev. Ed.) N.Y : Harcout, Brace, 1949.

25 – Nunnally, Jum C. Tests and Measurements Assessment and Perdiction. Mc. Grow Hill Book Compay Inc. New York 1959.

26 – Peatman, J.G. Descriptive and Sampling Statistices Harper Brothers 1947.

27 – Peters, C.C. and Van Voorhis, W.R. Statistical Procebures and their Mathematical Basis Mc–Graw Hill Book Co. 1940.

28 – Remmers, H.H. and Gage, N.L. and Runamel J.F. A Practical Introduction to Measurment and Evaluation. Harper Brothers, 1960.

29 – Shaffer, L.F. information which should be provided by test publishers and testing agencies on the Validity and

use of their tests : Personality tests. Proe, 1949 invit Conf. testing Service.

30— Skinner, C.E. Educational Psychlogy. Prentice Hall 1959.

31— Super, D. E. Appraising Vocational Fitness. Harper Brothers N.Y. 1949.

32— Torndike, R.L. and Hagen, A. Measurement and Evaluation in Psychlogy, and Eucation. John Willy, Sons Inc. 1957.

33— Travers, R. M. W. Eductional Measurement N.Y. Macmillan Co. 1955.

34— Walker, H. M. Elimentary statistical methods N.Y. Henry Holf & Co. 1947.

35— Walker, H. M. and Lev, J. Statistical Inference. Teachers Collage Columbia Univ., 1950.

36— Warren, H. C. Dictionary of Psychology Houghton Mifflin Co., 1934.

37— Wesman, A.G. Faking Personality test scores in a simulated employment situation. J. Appl. Psychol., 1952.

٣٩ – أحمد زكى صالح : علم النفس التربوى (الطبعة الخامسة)
القاهرة ، مكتبة النهضة ، ١٩٥٧

٤٠ – السيد محمد خيرى : الإحصاء فى البحوث النفسية والتربوية والاجتماعية
القاهرة ، دار الفكر العربى ، ١٩٥٦

٤١ – عطية محمود هنا : التوجيه التربوى والمهنى
القاهرة ، مكتبة النهضة المصرية ، ١٩٥٩

٤٢ – فؤاد البهى السيد : علم النفس الإحصائى وقياس العقل البشرى
القاهرة ، دار الفكر العربى ، ١٩٥٨

٤٣ – محمد عثمان نجاتى : علم النفس الصناعى
القاهرة ، مكتبة النهضة المصرية ، ١٩٦٠

٤٤ – نجيب اسكندر ولويس كامل ورشدى فام : الدراسة العلمية للسلوك الاجتماعى
القاهرة ، مؤسسة المطبوعات الحديثة ١٩٥٩

القسم الثاني

بناء المقاييس النفسية والتربوية ومميزاتها

الفصل الرابع

بناء المقاييس

- أهمية الإلمام بخصائص المقاييس
- تخطيط عام لمشروع بناء المقاييس
- خطوات إنشاء المقياس

أهمية الإلمام بخصائص المقاييس :

١ — إن كل مشتغل بمهنة يعرف الأدوات التى يستعين بها فى طرق العمل . يعرف خصائص كل أداة ويستطيع أن يختار الأداة المناسبة فى الوظيفة المناسبة . وهو يعرف ما هى الخصائص الرئيسية التى يجب أن تتوفر فى الأداة ومن ثم يستطيع أن يقيمها وأن يفاضل بينها وغيرها . والمشتغل بمهنة يعرف كيف يصون الأدوات التى يستخدمها ويعرف الخطوات المنظمة والاحتياطات الضرورية التى تجعله آمنا محافظاً على خلق المهنة فى نفس الوقت . فالطبيب مثلا يعرف الأداة التى يستعين بها فى قياس درجة الحرارة ، أو ضغط الدم ، أو قياس اضطراب وظيفة عضو معين . وهو فضلا عن هـــذا يعرف ، أو يجب أن يعرف ، أن بعض الأدوات يجب أن تكون من نوع معين عند استخدامها فى معالجة فرد يتميز بخصائص معينة . وهذا هو الحال تماماً بالنسبة للاخصائى النفسى سواء كان فى التربية أو الصناعة أو العيادة كما سنرى .

٢ — ولكن الطبيب لا يعرف كيف يصمم أى أداة من هذه الأدوات التى يستعملها ، وهو لا يقوم فعلا بهذا العمل ، وهذا لا ينتقص من قدره كطبيب ناجح أحياناً . فلا يقوم المشتغل بمهنة إذن ، فى العادة ، بتصميم أوثق الأدوات ارتباطا بمهنته . ولكنه يعرف ، بالتحديد ، وظيفة كل أداة ، ومتى يجب أن يستخدمها ، ومتى يحسن أن يستخدم أداة من نوع معين ، وكيف يفاضـــل بين عدة أدوات تؤدى نفس الوظيفة . بل ويعرف ، فضلا عن هذا ، ما هى المواصفات العامة التى راعاها من صمم هذه الأدوات . وأكثر من هذا فالمشتغل بمهنة يستطيع أن يحدد مدى ثقته بالحكم الذى بناه مستنداً إلى أدوات معينة :

٣ – والمشتغل بمهنة أيضاً يعرف كيف يخطط وحدة العمل التى سيقوم بها . وقد يخطط الأسس العريضة العامة التى سيتبعها فى حالات من نوع معين أو فى حالة بذاتها . فهو ، قبل معالجته لموضوع معين ، يحدد الأسلوب الفنى والطريقة المهنية التى ستجدى أكثر فى التعامل مع عميل أو مشكلة من نوع معين . وهو إلى هذا يتوقع أن يصل إلى أحكام تتعلق بجوانب المشكلة موضوع الدراسة . فهو يعرف نوع التقييم الذى سيقوم به أو يصل إليه عند اتباع أسلوب معين واختبار أداة معينة . وهو فى كل هذا ، عند ما يقوم بالعمل فعلا ، يوفر ظروفاً معينة ويتحاشى أخرى ويتحكم فى ثالثة : وباختصار فهو يعد نفسه ويعد أداته ويعد مشكلته ، يعدها جميعاً للبدء فى العمل . وهكذا الاخصائى النفسى .

٤ – والمشتغل بمهنة يعرف ، بكل تأكيد ، الوظائف التى يخدمها استخدامه لأداة معينة . فلكل أداة ، عنده ، وظائفها ، فما يقيس نسبة الرطوبة فى الجو ، وما يقيس الضغط الجوى عند ارتفاع معين ، يقيسها فعلا وبالتأكيد ولا يحتمل إطلاقاً أن يكون جهاز الضغط قد قاس كثافة المادة مثلا : وهو فى هذا يختلف عن الاخصائى النفسى الذى غالباً ما تكون ثقته ، محددة بنسبة معينة ، فى الوظائف التى تقيسها الأداة التى يستخدمها . بل ، وأحياناً ، يكتشف الأخصائى النفسى أن الأداة التى أراد بها قياس ظاهرة ما قاست أخرى بجوارها أو قاست أخرى ولم تقس ما أراد .

٥ – والمشتغل بمهنة يكرر استخدام نفس الأدوات فى مئات الحالات فتعطى له نتائج ثابتة أو متقاربة جداً . وهو فى هذا يختلف عن الاخصائى النفسى الذى لو كرر استخدام نفس الأدوات فى نفس الوظائف لوجد اختلافاً فى النتائج بين المرات المختلفة ، وإن كان هذا الاختلاف أحياناً كبيراً وأخرى مقبولا .

٦ — والمشتغل بمهنة يستطيع أن يحدد مدى تأكده من سلامة النتائج التى سيصل إليها . فالطبيب مثلا قبل إجراء العملية المعينة قد يعلن أن نسبة الخطورة فى العملية هى كذا فى المائة ، وأن النجاح محتمل بنسبة معينة ، وأن المضاعفات ستتعلق بجوانب معينة ، وأن إجراء العملية والعود إلى الحالة الطبيعية سيستغرق مدة تتراوح فى حدود معينة . وهو ، وإن كان قد حدد نسبة للثقة وفترة العلاج والجوانب التى ستتأثر بالعملية إلا أنه حدد هذا كله بنسبة مئوية أو فى مدى معين . ومع ذلك كثيراً ما تخرج النتائج الواقعة عن توقعاته . وهكذا الاخصائى النفسى ، تماماً ، يضع مستوى للثقة ، ويحدد جوهرية أو دلالة النتائج ، ويخطط البرنامج ، بل ويحدد الخطأ المحتمل والخطأ المعيارى . وبرغم هذا قد تحدث نتائج واقعة تخرج عن هذه التوقعات . وهذا لا يعيبه فليس الضبط هنا أمراً سهلاً ميسوراً وليست الحياة النفسية أرقاماً وتقديرات ولا حتى تقييمات ، فهى كل متفاعل متكامل ، إلى جوار ثباتها النسبى وخضوعها لقوانين السلوك العامة ، لها مرونتها وتميزها الفردى . ونحن لا نجد هذا بالطبع فى المادة التى يتعامل معها المهندس أو الكيميائى أو الطبيب فكلهم يتعاملون مع مواد ملموسة .

٧ — والطبيب ، لكى يصل إلى تشخيص صادق يستعين بمحكات أخرى ثبت صدقها أو محكات ترتبط تجريبياً بالوظائف التى يتناولها . فالصدق التلازمى عنده هو الاستعانة بفريق من الأطباء المبرزين فى حالات معينة (كونسولتو) . أى أنه يستعين بأحكامهم لمعرفة مدى صدق حكمه . وصدق الوقائع الخارجية عنده هو الاستعانة بتأثير أو مضاعفات علاجات معينة فإذا تحسن العضـــو أو الوظيفة التى يعالجها كان حكمه صادقاً . وهكذا الاخصائى النفسى يستعين بفريق Team يتناول معه الجوانب المختلفة كل فيما تخصص فيه .

٨ ـ والاخصائى النفسى لا يبنى حكماً ، أو هكذا يجب ، بناء على أدلة تبينها عند استخدام أداة واحدة مهما بلغ ثباتها وصدقها وشمولها . فهو يستعين بأدوات مختلفة ، ويجمع بيانات عن عدة أوجه من الحياة النفسية . والبيانات والأدوات كل واحدة تدعم الأخرى وتكشف عن صدقها وثباتها ، ومثله فى ذلك مثل الطبيب تماماً ، إذ يستعين الأخير بالتحاليل المختلفة ، والكشف بالأشعة ، والقياسات المتعددة ، كل يبين صدق الآخر ويكشف عن ثباته .

٩ ـ والاخصائى النفسى عند دراسة الفرد أو الجماعة ، يبدأ بتكوين فكرة كلية تقريبية ، ثم يشرع فى تحليل هذا الكل أو الجشطلت إلى عناصره ويدرس كلا منها . ثم يعيد بناء ما حلله ليكون صورة كلية أدق من الأولى . فهو يطبق اختبارات المسح ، ليكون فكرة عامة عن الفرد أو الجماعة ، ثم يقيم ، ثم يضع علامات الاستفهام على نقاط الضعف والقوى فى التكوين النفسى لمن يدرسه . ويبدأ فى دراسة أكثر تفصيلا لجوانب الإشكال . وبعد هذا يدرس ديناميات الحالة بطرق مثل الطرق الإسقاطية وأخيراً يكون صورة كلية متكاملة يعرف تفاصيلها وإبعادها ونسبها ، وماهى الزواية التى ينظر منها إلى هذه الصورة وتلك الأبعاد .

١٠ ـ والطبيب والفلكى والمشتغل بالقياسات الألكترونية ، كل منهم يضع مستوى من الدقة إذا وصل إليه القياس كان قياساً صادقاً وثابتاً . وهم فى هذا يحددون ، إن تفصيلا وإن ضمنا ، نسبة ثبات وصدق قياساتهم فإذا لم يتجاوز القياس حد الخطأ المعيارى هذا ، كان القياس والمقياس صالحين . هذا على حين أننا نجد أن الاخصائيين النفسيين لم يتفقوا على مستويات يحددون بها صدق وثبات مقاييسهم .

١١ – عالم الطبيعة وعالم الكيمياء يحللان مواداً يعرفان درجة نقائها . بل ويستطيع كل منهما أن يصل إلى عناصر نقية يقيس فاعليتها . أما المشتغل بالقياس النفسى والتربوى فقد كان ، إلى حين قريب ، يقيس عوامل مستقلة ليبين علاقتها بعوامل متغيرة فيجد أنه لا يقيس هذه ولا تلك بالتحديد ؛ وإنما يقيس عوامل متدخلة أو متوسطة ؛ وقد بدأت الطرق الفنية للوصول إلى عوامل ذات نقاء أكثر ، بدأت تتقدم حثيثاً بفضل عدة مناهج علمية أهمها التحليل العاملى .

١٢ – وأخيراً نجد أن كل مشتغل بمهنة يعرف ، بدقة ، شروط إجراء عملياته وظروفها ولديه المعايير أى المستويات التى يرجع إليها عند الحكم ؛ فقد مرت الأدوات والطرق الفنية فى دور التجربة بعملية التقنين ، وأصبح ، بالتالى ، استخدامه للأداة أو الطريقة قانونياً أو مشروعاً . والأخصائى النفسى بالمثل يجرب أدواته على عينة تسمى بعينة التقنين ، ويصل من ذلك إلى تحديد الشروط والظروف والمعايير والصدق والثبات ؛

وهكذا رأينا أن الأخصائى النفسى يحتاج إلى تحديد الوظيفة التى يقيسها بتحليلها ، وإلى تحديد الأغراض التى تخدمها أدوات معينة بتحليلها ، ثم يطابق بين هذه وتلك كى يختار الأدوات المناسبة . كما رأينا أنه يختار ويقيم ويحكم ويفاضل بين الأدوات وإن كان لا يقوم بتصميمها ، وهو فى هذا يتساوى مع غيره من الأخصائيين لا يلزم أن يكون مصمم الأداة التى يستخدمها . إذ أن تصميم الاختبارات ليس بالعمل السهل كما سنرى . والمشتغل بالقياس النفسى عليه أن يضع تخطيطاً مستنداً إلى إعداده العلمى ، ومعتمداً على دراسته للمشكلة ، ومتناسباً مع إمكانيات القياس لديه .

والأخصائى النفسى عليه ، كما رأينا أن يتأكد من صدق قياس الأداة لما زعمت أنها تقيسه . كما يلزمه أن يحدد مدى ثبات النتائج التى وصل

إليها . بل وعليه أن يحدد مدى صدق وثبات النتائج التى قد يصل إليها ونوع هذه النتائج ، وأن يهتم بنوع من المعلومات أكثر من غيره أهمية فى تأثيره على ما يقيس ، فيحدد بذلك نوع المثيرات التى سيقدمها وطريقة تقديمها أى مسألة صياغتها فى فقرات .

ورأينا أيضاً أن الحياة النفسية لا تلخص فى صفحة نفسية تعبر عن مستويات كمية . وإن كان هذا أمر يعين الدراسة ، إلا أنه غير كاف إذ يجب أن يقوم بالتحليل الكيفى ويدرس الديناميات والعلاقات الوظيفية بين المتغيرات . وهو فى دراسته هذه ، كغيره من المهنيين ، يعتمد على آراء غيره ممن يكونون معه فريقاً متكاملا . ويلجأ إلى الإفادة من أدوات مختلفة كل يقيم جانباً أكثر مما يقيم جوانب أخرى . وهو فى كل هذا يكون صورة كلية انطباعية ثم يحلل هذا الكل حتى يصل إلى تكوين صورة كلية أدق من الأولى . على أن هذه الصورة لا توضع بطريقة حتمية مؤكدة تماماً إذ يلزم اعتبار تدخل عوامل لم يستطع ضبطها . وفى النهاية يلزم الأخصائى النفسى أن يعرف حدوده وحدود الأدوات التى يستخدمها وظروف القياس المختلفة .

والأخصائى النفسى وإن كان لا يلزم أن يصمم الاختبارات بنفسه كما رأينا ، يحتاج لكى يعرف الوظائف والمعايير وشروط الإجراء وطريقة التصحيح ومعنى الدرجة ودلالتها والتفسير الكيفى وثبات وصدق الأداة ودرجة ثقته فيها ، يحتاج إلى أن يعرف كيف صممت الأدوات التى يستخدمها . ونحن لا ننكر ، بالطبع ، على الأخصائى النفسى القيام بتصميم الاختبارات وإلا فمن يصممها ؟ هذا برغم ما يكتنف هذا العمل الضخم من مشقات وما يلزمه من احتياطات وما يقتضيه من علم نظرى وعملى . إلا أن الأخصائى النفسى قد يجد نفسه أمام مشكلة لم تعد تجدى

فيها اختبارات موجودة ، كما قد يجد أن الاختبارات التى تقيس ما يريد قياسه قننت على عينات من نوع يخالف طبيعة المجتمع الذى يريد دراسته . وهنا يلزمه أن يقوم بتصميم اختبارات جديدة إن كان له ذلك .

تخطيط عام لمشروع بناء المقياس

أولا ـ المشكلة :

برغم ما يكتنف إعداد المقياس من مشقة جمة إلا أن الأخصائى النفسى قد يجد نفسه إزاء موقف يلزمه بأن يقوم بهذا العمل أو أن يسهم فيه . فالأخصائى النفسى قد يقوم بعمليات الاختيار التربوى أو المهنى ، أو التوجيه التربوى أو المهنى ، أو الإرشاد النفسى ، أو عمل يدخل فى نطاق الخدمات النفسية الفردية أو الجماعية ، وهو فى كل هذا يحتاج إلى أداة موضوعية مقننة لتقدير الأفراد الذين يتعامل معهم بطبيعة عمله . وهو قد يجد تلك الأداة فى مقاييس جيدة تناسب الوظيفة التى استخدمها ، إلا أن مثل هذه المقاييس ، لسعة انتشارها ، قد تكون فقراتها معروفة . وهنا يفقد الاختبار الموجود قيمته فلا يصبح مقياساً للقدرة أو السمة ، بل اختباراً للتذكر والاسترجاع . فيجد أن الأفراد يعرفون الإجابات الصحيحة غالباً ، ومن ثم لا يستطيع أن يحدد من هو أصلحهم .

والأخصائى النفسى قد يجد المقياس الذى ينشده ، وهو يتعرف عليه غالباً من اسمه ، ولكن هذا المقياس أحياناً يقيس شيئاً آخر . وأحياناً أخرى لا يمثل المقياس الوظيفة المراد قياسها فحسب ، بل وأيضاً وظائف أخرى قد تكون غير ذات دلالة بالنسبة لمهنة أو دراسة معينة . وقد يلائم اختبار موجود بالفعل الغرض المنشود ، ولكنه كثيراً ما يكون قد جرب

وقنن واشتقت معاييره من جماعات تختلف فى طبيعتها عن الجماعات المراد دراستها . ومن هنا يلزم إعــداد اختبار يقنن على الجماعات التى سيطبق عليها :

والأخصائى النفسى ، فى كثير من الأحيان ، يجد أن الاختبار الموجود يقيس وظيفة يتطلبها عمل معين . ولكنه لا يقيس كل أو أهم الوظائف التى يقتضيها النجاح فى هذا العمل . ولهذا فهو يقوم بإعداد اختبار أو سلسلة من الاختبارات تغطى هذه الوظائف معاً . وقد يحدث أن يجد الأخصائى اختباراً يقيس وظائف معينة ولكنه لا يرقى إلى تقييم الأفراد الذين وصلوا إلى مستوى علمى أو مهنى معين . كما أن الاختبار قد يقيس الوظيفة ولكنه يفعل هذا من وجهة نظر معينة . بينما يريد الأخصائى تقييم نفس الوظيفة من زاوية أخرى . وهنا يصمم اختباراً يصحح من وجهــة نظر معينة .

وقد تكون هناك صعوبات عملية تحول دون استخدام اختبار موجود بالفعل . فالأخصائى النفسى محدود ، فى البرنامج الذى يستخدم فيه الاختبارات ، بتوقيت معين يخطط على أساسه الزمن اللازم لإجراء الاختبار . فإذا كان الاختبار الموجود يتطلب وقتاً يزيد على ما يسمح به التوقيت ، كان استخدام الاختبار هنا مستحيلا ، ولزم إعداد اختبار يفى بالغرض فى حدود الزمن المحدد . كما أن بعض الاختبارات الموجودة تتطلب من القائم بها خبرة وإعداداً فنياً غير متوفر فيمن سيقومون بإجرائها فى مؤسسة أو مدرسة معينة . ولذا يلزم إعداد اختبار يسهل على الأخصائيين إجراؤه . ثم إن الاختبار الموجود قد يتطلب فى إجرائه نفقات لا تحتملها ميزانية المؤسسة أو المدرسة التى يعمل بها الإخصائى ، وهنا يلزم إعداد اختبار أكثر اقتصاداً . وبالمثل قد تكون مواد الاختبار من نوع يتلف

عندما ينقل من جهة إلى أخرى أو يصعب نقلها بطريقة عملية ، وهنا يستعاض عنها بمواد يمكن نقلها بطريقة عملية .

وأخيراً فإن الأخصائى قد يجد الاختبار الملائم الذى لا يعيبه أى نقص مما سبق ذكره . ولكنه يعطى الاختبار لأخصائيين مختلفين فيطبقونه على بعض الأفراد ويكتشف أن درجة الفرد الواحد تختلف باختلاف من قام بإجراء الاختبار عليه ، فالاختبار إذن ذاتى . وقد يقوم أخصائى واحد بإجراء الاختبار عدة مرات على نفس الفرد دون أن يحدث تغير جوهرى فى الفرد ، وبرغم هذا تختلف درجة الفرد من مرة إلى أخرى . وكأن الاختبار لا يعطى نتائج ثابتة . والعيبان ، الذاتية وعدم ثبات النتائج ، يجعلان الاختبار الموجود أداة لا يمكن الاعتماد عليها . ومن ثم يصمم الأخصائى اختبارا يمكن الاعتماد عليه من هذه الناحية .

ولا يلزم فى كل الأحوال التى سبق ذكرها ، أن يقوم الأخصائى بتأليف الاختبار من أوله إلى آخره ، فهو قد يكتفى بتعديل اختبار موجود حتى يتلاءم مع هدف منشود . وأيا كان الجهد المبذول ، سواء فى التأليف أو التعديل ، فهو جهد يهدف إلى إعداد اختبار جيد . والاختبار الجيد لا يوجه إليه نقد من الانتقادات السابقة . فهو اختبار يمثل وظيفة . وهى وظيفة نقصد قياسها لأنها تتعلق بالنجاح فى عمل معين . وهو عمل يشمل جوانب فى الفرد يمثلها الاختبار تمثيلا صادقا ويعطى عنها درجة . وهى درجة لها عندنا معايير ، تمكننا من تفسيرها . وهو تفسير ثابت لاختبار صادق فيما يريد قياسه : وهو قياس عملى من حيث الوقت والتكاليف وسهولة النقل وتوفر من يستطيعون إجراءه .

ثانيا ــ تحديد الأهداف :

عندما تتضح المشكلة بالطريقة السابقة يكون الاخصائى قد اقترب من الوصول إلى تحديد أهداف إنشاء أو إعداد اختبار جديد : فيلاحظ الاخصائى مثلا أنه يجب أن يصمم مقياسا يعينه فى اختيار أفضل المتقدمين للعمل أو الدراسة . ويكون الهدف هنا هو الاختيار وتكون المشكلة هنا هى أن عدد المتقدمين للاختيار أكثر من عدد الوظائف الشاغرة الجديدة . وقد يكون الهدف هو إعداد أدوات تساعد فى التوجيه المهنى أو التربوى وتكون المشكلة هنا هى أن أمام الفرد عدة فرص قد ينجح فى أحدها ويتعين تحديد هــذه الفرصة بالذات . وفى حالتى التوجيه المهنى والتربوى والاختيار المهنى والتربوى يهــدف المقياس إلى مسح وتقييم الفرد .

وقد يهدف الأخصائى النفسى إلى تحديد نقاط الضعف والقوى فى أفراد يدرسهم أى إلى تشخيصهم ويكون هدف الاختبار إذن هو التشخيص وتكون المشكلة هى التقييم الكيفى عن طريق استجابة المفحوص لموقف الاختبار . والأخصائى قد يريد أن يتنبأ بالنجاح أو الفشل الذى سيحققه من يلتحق بمهنة أو دراسة . وقد يهدف الأخصائى إلى معرفة ما يتوقع من تحصيل المفحوص لأهداف تالية بناءاً على ما حصله من أهداف سابقة . وقد يريد الأخصائى معرفة ما إذا كانت حالة مفحوص معين تنبئ بتخففه من أعراض مرضية معينة بعد اتباع أسلوب تشخيصى وعلاجى معين . وفى كل هذه الحالات يكون هدف الاختبار هو التنبؤ . وتكون المشكلة هى معرفة مدى ارتباط السلوك الراهن بالسلوك المقبل أى معرفة دلالة نتيجة الاختبار على سلوك المفحوص مستقبلا .

والأخصائى ، فضلا عن كل هذا ، قد يريد تصنيف جماعة كبيرة إلى مجموعات متناسقة صغيرة تتلقى تدريباً مهنياً معيناً أو منهجاً دراسياً متقدماً أو تلقى عناية أكبر فى دراسة معينة . وهناك أهداف أخرى كثيرة قد يحدد الأخصائى نفسه بها فى إنشاء المقياس كتقييم العمل أو ترشيح من يستحقون الترقية أو العلاوة أو . . الخ . وأخيراً وليس آخراً قد يكون الهدف هو خدمة البحث العلمى فى ميادين علم النفس والتربية :

ثالثاً — الإعداد للمشروع :

١ — بعد وضوح المشكلة وتحديد الهدف يقوم الأخصائى بعرض فكرة المشروع على الخبراء فى القياس . ويناقش معهم أهمية المشكلة وقيمة الأهداف وما سيحققه الاختبار . محاولا التحقق من صحة الخطة أو الحصول على توجيهاتهم قبل المضى فى المشروع .

٢ — يحصل الأخصائى على تصريح يطلب فيه المسؤولون من مختلف المدارس أو الهيئات تقديم العون للأخصائى لإجراء تجاربه ودراساته على أن لا يكون لذلك تأثير فى سير العمل أو الدراسة .

٣ — يقوم الاخصائى بتقـدير تكاليف مشروع الاختبار ويقدم تقريراً بذلك إلى المسؤلين إذا كان يبغى إنفاقهم على المشروع أو الحصول على مساعدة مالية لتحقيق المشروع .

٤ — يعين المسؤولون هيئة أو فرداً يشرف على المشروع ، وهـذا المشرف يجب أن يكون خبيراً وعالماً فى الوقت نفسه . وهو غالباً ما يكون أستاذاً بالجامعة .

٥ — يقوم المشرف باختيار فريق من الأخصائيين يساعده فى المشروع ويدرب هذا الفريق على عمليات القياس المختلفة شارحاً لهم طبيعة وأهداف المقياس الجديد :

٦ – يضع المشرف والفريق معاً تخطيطاً بالحدود الزمنية للمشروع ويقدمونه للمسؤولين الذين يلائمونه مع اعتبارات أخرى هامة فى نظرهم .

٧ – يقوم المشرف والفريق بالاطلاع على الاختبارات ومراجع القياس الموجودة حتى يكونوا محيطين بالمشكلة ، ويناقشون الاخصائى فيها .

٨ – يضع المشرف تخطيطاً لتوزيع أدوار وتحديد مسئوليات أعضاء الفريق .

٩ – يحدد المشرف والفريق الأعمال أو الدراسات التى سيطبق الاختبار عليها .

١٠ – يجمع الفريق معلومات وملاحظات عن هـذه الأعمال أو الدراسات .

١١ – يقوم الفريق بتحليل هذه الأعمال أو الدراسات .

رابعا – تحليل المضمون :

نقصد بتحليل المضمون Content Analysis تحليل العمل Job Analysis إن كنا بصدد مهن أو تحليل برنامج الدراسة Curriculum Analysis إن كنا بصدد الدراسة . ولا نقصد هنا تعريف هذه المصطلحات التى سبق أن عرّفناها ، بل معرفة ماذا تفيدنا فيه هذه الخطوة وما هى أهميتها . فلا شك فى أن الاختبار ، لكى يكون مبنياً على بيانات واقعية ، يجب أن يؤسس بعد تحليل العمل . وتحليل العمل يعرفنا بالوظائف التى ترتبط بالنجاح فى العمل ومدى أهمية كل وظيفة فى هذا العمل . أما تحليل المنهج الدراسى فهو يعرفنا بأهداف التربية وأهداف تدريس منهج بذاته . وبناءاً على هذه الأهداف يتبنى المنهج أو يتطلب وظائف معينة . هذه الوظائف يجب أن نحددها

فى تحليل المنهج . وكأن تحليل المضمون ، سواء كان عملا أو منهجاً دراسياً ، هو « دراسة علمية شاملة دقيقة تعتمد على عدة طرق ومصادر ، وتهدف إلى معرفة واجبات ومقتضيات المهنة أو الدراسة ومسوئلياتها ، وتحديد ما يجب توفره فى القائم بها » .

وهكذا نرى أنه يمكننا أن نتحدث فى تحليل المضمون عن :

(أ) تحليل المنهج الدراسى عند تصميم اختبارات التحصيل .

(ب) تحليل العمل عند تصميم الاختبارات النفسية فى المهن .

(أ) تحليل المنهج الدراسى : يتصفح المدرس الكتب المقررة ، ويراجع مذكرات الدروس التى ألقاها ، عندما يكون بصدد وضع امتحان نهاية العام . ولكنه إذا اعتمد على الكتب والمذكرات فقط لما تبين مدى صدق امتحانه . وهو قد يسائل نفسه هنا قائلا : « أى هذه الموضوعات يصلح موضعاً للسؤال ؟ » . ولكنه إذا كان قد حدد أهدافاً للتربية بوجه عام وللمدرسة بوجه خاص ، بل وحدد أهداف تدريس منهج دراسى معين ، لأصبح فى موقف أفضل . وهو هنا يسائل نفسه الأسئلة الآتية :

١ — ما الذى أريد أن يفعله التلاميذ بالحقائق التى تعلموها ؟

٢ — ما الذى سيحتاجه التلميذ ، فيما بعد ، فى الدراسة أو العمل ؟

٣ — ما هى العلاقات التى يجب أن يخرج بها التلميذ من المفاهيم التى شرحت له ؟

٤ — ما هى الآراء والأفكار التى يجب أن يضيفها التلميذ من واقع خبرته ؟

٥ — ما هى المهارات التى يجب أن يكون التلميذ قد تعلمها ؟

وتعطى للموضوعات المتعددة وتفاصيلها أهمية نسبية فى محتوى المنهج ،

وبالتالى فى محتوى الاختبار ، على أن يشملها الاختبار فى حدود الزمن المسموح به . وبدلا من وزن الدرجات حسب أهمية الأسئلة بالنسبة للأهداف ، يمكن وضع عدد من الأسئلة يتناسب مع جوهرية كل موضوع . ويقرر مصمم الاختبار أى نوع من الفقرات يناسب غرضه ؛ فقرات المقال ، أو التكميل ، أو المزاوجة ، . . الخ . ويستطيع المصمم أن يقسم الاختبار إلى عدة جوانب أساسية وتحت كل جانب يحدد نوع الأسئلة الملائمة والعدد المقترح منها . على أن كل هـــذا التحديد مبدئى وغالباً ما يتغير تبعاً لما تكشف عنه البحوث ، ولكن هذا لا يقلل من قيمة عمل خطة واضحة .

(ب) تحليل العمل : هو الدراســـة العلمية المنظمة للعمل لتحديده ومعرفة طبيعته وعملياته وأهميـــة كل عمليـــة منها ومعرفة ما هى الخصائص والشروط التى يجب أن تتوفر فيمن يقوم به ، وحصر الواجبات الخاصة بكل وظيفة . ويقول سوبر D. Super « يقدم تحليل العمل لمن يصمم الاختبار قائمة بالاستعدادات والسمات التى تعتبر هامة فى ذلك العمل . ولكن هذه القائمة تخضع لعيبين خطيرين : وهما ذاتية الأدلة وعدم التأكد من أن عاملا Factor معيناً يميز هذا العمل عن غيره ، حتى لو برهن العامل على أهميته » :

وتحليل العمل فى خطواته النهائية يصل إلى بيانات تلخص فى قائمة تسمى وصف العمل Job Description تحدد العمليات المتضمنة فى هذا العمل وطبيعته وأهمية كل منها وكيف تؤدى ، أى أنه يجيب عن الأسئلة الثلاث الآتية :

١ ــ ما هى العمليات والواجبات المتضمنة وما طبيعتها what the worker does ?

٢ — لماذا تُؤدى كل عملية منها وما هى الأهمية النسبية لكل عملية منها
? Why he does it

٣ — كيف تُؤدى هذه العمليات وما هى الطرق العامة المتبعة فى كل منها ? how he does it

كان هذا فيما يتعلق بوصف العمل . أما فيما يتعلق بوصف العامل فان الجزء الخاص به يلخص فيما يسمى بمقتضيات العمل Job Requirements أو مواصفات العمل Job Specifications . وقائمة شروط العمل هذه تحدد الشروط والخصائص العقلية والبدنية والمؤهلات والخبرة الواجب توفرها فيمن يقوم بهذا العمل . أى أنها تجيب عن السؤال الآتى :

٤ — ما هى المهارة المتضمنة The Skill involved ?

وهكذا رأينا أن تحليل المحتوى ، والذى يشمل تحليل المنهج الدراسى كما يشمل تحليل العمل ، يمدنا من ناحية بخصائص العمل ومن أخرى بالخصائص التى يجب توفرها فيمن يقوم به . ومن ثم نستطيع أن نحدد : ما هى المجالات التى ستغطيها أسئلة الاختبار ؟ وما هى الأهمية النسبية لكل مجال منها ؟ وما نوع الأسئلة الذى يناسب كل مجال ؟ وما عدد هذه الأسئلة لكل ؟ فإذا حددنا هذا أمكننا أن نحدد طبيعة المجتمع الذى سيجرب عليه الاختبار ثم يقنن ثم يستخدم . وحددنا المعايير وطريقة تصحيح وتفسير كل سؤال .

خامسا — مواصفات الاختبار :

إلى هنا تكون صورة غامضة قد تكونت لدى مصمم الاختبار . صورة ناقصة تقريبية عما سيكون عليه الاختبار أخيرا . هذه الصورة تنمو وتتضح كلما تقدم المصمم فى تخطيطه ، ولكن معالمها لا تتضح إلا عند ما يبدأ فى إنشاء الاختبار فعلا . وقبل هذا الإنشاء على المصمم أن يحدد نفسه

كتابة بمواصفات الاختبار الذى ينوى إنشاءه ، وهذا التحديد سيوضح له ولمساعديه ما يحاول أن يعمله . وهـــذه المواصفات يجب أن تغطى النقاط التالية :

١ — الغرض : وصف دقيق للغرض الرئيسى للاختبار والأغراض الأخرى .
٢ — ماذا يقيس : وصف دقيق للخاصية أو الوظيفة التى يقصد قياسها ، مع ذكر الظروف والعوامل التى قد تؤثر فى نتائج القياس .
٣ — وصف الاختبار : مادة الاختبار ، عدد الأجزاء ونوع وعدد الفقرات فى كل جزء ، تعليمات الإجراء ، طريقة التصحيح .
٤ — العينة : تحديد الجماعات التى يجرى عليها الاختبار ، تحديد حجم وطبيعة العينة التى تحسب منها خصائص الاختبار والمعايير .
٥ — منهج البحث : خطوات البحث ، الطرق الإحصائية لاختيار الفقرات وترتيبها وحساب خصائص الاختبار والمعايير .

على أن كل هذه الخصائص تظل انسيابية فى المراحل الأولى من إنشاء الاختبار ، بمعنى أنها تتطور وتتحدد أثناء عمليات التقنين التالية . ووصف الاختبار إذا غطى النقاط السابقة كان ممثلا لطريقة التفكير فى موضوع الاختبار وقت إنشائه . وإن كان يجب أن يوضع بطريقة تقبل التعديل والمراجعـــة على أساس الأفكار الحديثة المتعلقة بالموضوع عموماً وفى ضوء البيانات الواقعية المشتقة من تجريب عينة من التعليمات والفقرات ، أى الصورة الأولية للاختبار خصوصاً . وهكذا تنمو المواصـــفات بنمو الاختبار .

ومواصفات الاختبار ، إذا وضعت هكذا ، تكون ذات قيمة فى

بلورة خطة الاختبار فى ذهن مصممه . كما أن لها قيمة فى توضيح الخطة الظاهرة لمعاونى مصمم الاختبار ، بحيث يعملون فيه بأعلى كفاية فتثار وتوجه الاقتراحات الإيجابية وأوجه النقد الموضوعية . وتنحصر الفائدة الثالثة لتعيين خصائص الاختبار فى كونها مرشداً يعين فى تأليف صور مكافئة من الاختبار نفسه . فإذا صيغت عبارات واضحة عما يتضمنه الاختبار من حيث نوع ومحتوى وصعوبة الفقرات ، وإذا عينت العبارات الخطوات العملية لتأليف واختيار الفقرات ، تيسر إعداد صور تالية من الاختبار تكون مكافئة حقاً للصورة الأولى فى الوظائف المقاسة والمدى المشمول ومستوى الصعوبة .

خطوات إنشاء المقياس :

بعد أن تكون المشكلة قد اتضحت فى ذهن واضعى الاختبار ، وبعد أن دفعتهم الحاجة إلى إعداده ، تحدد أمامهم الأهداف . وبعد أن تتضح المشكلة ويتحدد الهدف يقومون بالإعداد للمشروع ، وهنا يحللون العمل أو منهج الدراسة (حسب نوع الاختبار) ، ومن ثم يحددون المواصفات التى سيكون عليها الاختبار . وهنا يبدأ الفريق فى الخطوات العملية لإنشاء الاختبار ، وهى تنحصر غالباً فيما يلى :

١ — اختيار نوع وحدات الاختبار ؛

٢ — اختيار نوع مواد الاختبار.

٣ — تأليف الفقرات ؛

٤ — تجربة الفقرات ؛

٥ — تحديد مستوى صعوبتها ؛

٦ — البحث عن المحكات ؛

٧ – جمع بيانات عن المحكات .

٨ – تحديد درجات الأفراد على المحكات .

٩ – مقارنة درجات الأفراد على الفقرات بدرجاتهم على المحكات .

١٠ – تحديد طريقة التصحيح .

١١ – تعيين أوزان الفقرات .

١٢ – تحديد نوع الجماعات التى سيطبق عليها الاختبار .

١٣ – تحديد طرق وعمليات الإجراء .

١٤ – اختيار عينة التقنين .

١٥ – إجراء الاختبار على عينة التقنين .

١٦ – اشتقاق المعايير .

١٧ – تحديد الدرجات الفاصلة .

١٨ – حساب الثبات والصدق .

١٩ – التأكد من صلاحية الاختبار فى قياس ما أريد به على المجتمع الأصلى .

٢٠ – نشر الاختبار وتحديداته .

٢١ – شروط إعداد الأخصائيين الذين سيستخدمون الاختبار .

٢٢ – متابعة إجراء بحوث على الاختبار ونشرها .

وهكذا نرى أن أهم الخطوات العملية فى إعداد الاختبار تنحصر فى :

(ا) إعداد الصورة الأولى للاختبار .

(ب) تجريبها .

(حـ) تعديلها .

(د) الإجراء المبدئى .

(هـ) عمل صورة أو صور مكافئة من نفس الاختبار .

(و) التقنين .

(ز) الصورة النهائية من الاختبار .

(ح) نشر كراسة التعليمات ونماذج التصحيح والدراسات .

لم نرد الاستطراد فى كل نقطة من النقاط السابقة إذ ستكون موضعا للبحث فى فصول تالية ٥

الفصل الخامس

إعداد فقرات الاختبار

- مقدمة .
- ماذا يجب .
- كيف نصمم الفقرات .
- أنواع فقرات الاختبار .
- مقارنة الأنواع .
- قواعد عامة فى إعداد الفقرات .
- التحليل الإحصائى للفقرات .
- تصحيح الفقرات .

مقـــدمة :

يتكون الاختبار عموماً من عدد من الفقرات : والفقرة item هى وحدة فى الاختبار : هذه الوحدة قد تكون سؤالا واحداً بسيطاً أو أكثر . كأن يوضع سؤال يحدد المطلوب فى الإجابة عن عدة أسئلة تالية : والفقرة تسأل عن إحدى العمليات أو المتغيرات النفسية ، وتكون لها مادة من نوع معين : وتصحح الفقرات بطرق مختلفة .

وهكذا نرى أنه يجب :

١ ــ تحديد النوع المفضل من الفقرات .

٢ ــ اختيار أنسب أنواع المواد :

٣ ــ تحديد المستوى السائد أى ما هو النوع من المضمون الذى سيغلب على بقية الأنواع .

٤ ــ تعيين أوزان الفقرات .

٥ ــ اختيار أنسب طريقة للتصحيح .

ونحن نجد أننا فى هذه الأمور لسنا بصدد قواعد عامة جامدة تطبق فى جميع الأحوال . فإن اختيار نوع الفقرات ، والمواد ، والمحتوى السائد ، والأوزان ، وطريقة التصحح ، كل هذا يتحدد تبعاً لعدة اعتبارات أهمها :

(ا) الوظيفة أو الوظائف التى يقيسها الاختبار .

(ب) الغرض الذى سيستخدم الاختبار فيه .

(حـ) نوع المجتمع الأصـــلى المراد قياسه وخصائصه ومستواه فى الوظائف المقاسة .

(د) إمكانيات الأخصائيين الذين سيصممون الاختبار ومن سيجرونه :

(هـ) التكوين العلمى لمصمم الاختبار وما يراه أساسياً :

(و) حدود الوقت والتكاليف .

(ز) مستوى الثقة الذى يراه مصمم الاختبار كافياً إذا توفر ، من حيث ثبات وصدق الاختبار إذ أنهما يتأثران بالفقرات كما سنرى .

وبالرغم من كل هذا فمن المؤكد أن الاختبار ، باعتباره أداة موضوعية مقننة لقياس جانب أو جوانب من السلوك بطريقة ثابتة ، يجب أن يمثل الجوانب المراد قياسها . وهذا التمثيل لكي يكون كافياً يجب أن تتوفر فيه صفتي الكم والكيف . بمعنى أنه ما دام الاختبار عينة من السلوك ، فيجب أن يمثل أنواع ومستويات الجوانب المقاسة فيتضمن كل الأنواع والمستويات ويجب أن يمثلها من حيث أهمية كل نوع وكل مستوى فيتضمن فقرات يتناسب عددها مع أهمية الجانب الذى تمثله فى الوظيفة المقاسة . ولنفرض أننا بصدد عمل اختبار تحصيلى فى الحساب ، كيف نصمم فقراته ؟

أولا : نحلل محتوى منهج الحساب الذى درسته الفرقة الدراسية التى تصمم الاختبار لقياس تحصيلها .

ثانياً : نحـدد العمليات التى درست ولتكن الجمع ، الطرح ، الضرب ، والقسمة .

ثالثاً : نحدد نسبة كل عملية منها وأهميتها فى أهداف تدريس المادة كما درست بالفعل ولتكن هذه النسب هى : ٢٠٪ جمع ، ٢٠٪ طرح ، ٣٠٪ ضرب ٣٠٪ قسمة .

رابعاً : نمثل هـذه الأنواع وتلك النسب فى الفقرات بإحدى الطريقتين الآتيتين :

(ا) لنفرض أن عدد الأسئلة هو ٢٠ فيكون عدد الأسئلة موزعاً هكذا ، ٤ جمع ، ٤ طرح ، ٦ ضرب ، ٦ قسمة ولكل سـؤال درجة واحدة .

(ب) يمكننا أن نوزع عدد الأسئلة بالتساوى بين العمليات الأربعة فيخص كل من الجمع والطرح والضرب والقسمة ، ٤ فقرات أو أسئلة وهنا نعطى درجة واحدة عن الإجابة الصحيحة عن أى من أسئلة الجمع والطرح ، ودرجة ونصف عن الإجابة الصحيحة عن أى من أسئلة الضرب والقسمة .

وهكذا رأينا أننا حللنا المنهج الدراسى واخترنا مواد تناسبه وهى عمليات الحساب ، وحددنا ما تقيسه الفقرات وعددها وتوزيعها وأوزانها . حدث كل هذا ، بالطبع ، تبعاً لمستوى الفرقة الدراسية التى نقيسها فلا يصح مثلا أن نعطى عمليات قسمة مطولة لمن تعلموا القسمة البسيطة فقط فإن مستوى الصعوبة يجب أن يتناسب مع طبيعة المجموعة . بقى شىء آخر وهو تحديد نوع الأسئلة ، فقد يفضل المدرس إعطاء أسئلة يطلب من التلاميذ الإجابة عنها . وقد يعطيهم السؤال والإجابة ويتطلب من التلميذ أن يضع علامة صح أو خطأ على الإجابة . وقد يضع المدرس السؤال متبوعاً بعدة إجابات يختار التلميذ إحداها . وقد يضع المدرس السؤال فى شكل عبارات يترك خلالها مسافات يجيب فيها التلميذ بالتكميل . وقد يضع عدداً من الأسئلة فى عمود ويضع إجابات هذه الأسئلة فى العمود المقابل ، بعد أن يخلط الإجابات خلطاً عشوائياً ، ثم يطلب من التلميذ أن يكتب أمام السؤال الحرف الموجود أمام الإجابة .

وهكذا نجد أن هناك أنواعاً متعددة من الفقرات يختلف واضعوا الاختبارات فى تفضيلهم لأحدها على الآخر . ولكن يجب أن نراعى أن تصاغ الفقرات بحسب الطريقة التى درس بها التلاميذ المادة على أنه من المؤكد أن الاختبار يكون أفضل عندما يستخدم عدداً كبيراً من الفقرات ، وعندما تتنوع المواد المستخدمة . فما لا شك فيه أن الاختبار الذى يشمل عدداً كبيراً من الفقرات هو اختبار أفضل عموماً من الاختبار القصير . وليست هذه المسألة أيضاً مطلقة فى كل الحالات ، فأحياناً لا يكون هناك

متسع من الوقت لإجراء اختبار طويل . كما أن الاختبار إذا كان طويلا جداً تأثرت الإجابة عن أجزائه الأخيرة بعوامل التعب والملل ، وإذا كان على درجة كبيرة من القصر يزيد احتمال تدخـــل عوامل التذكر والحفظ ، على أنه بالرغم من عدم وجود قواعد جامدة فى وضع فقرات الاختبار ، هناك ملاحظات يجب اعتبارها — سنتحدث قريباً عنها .

أنواع فقرات الاختبار

١ — فقرات الصواب — الخطأ True-False items : وهى فقرات نقدم فيها للمفحوص جملة أو عبارة أو مشكلة أو مسألة ونطلب منه أن يحدد صحة الفقرة بوضع علامة ✓ أو خطأ (×) أو وضع خط تحت كلمة صحيح أو خطأ .

٢ — فقرات التكميل Completion items وهى فقرات نطلب فيها من المفحوص أن يكمل عبارة أو مسألة بإضافة العنصر المناسب مثل : حاصل ضرب ٥ × ٦ هو . . .

٣ — فقرات الاختيار من عدة إجابات muliple-choice items وهنا نعطى المفحوص السؤال ونتبعه بعدة إجابات إحداها صحيحة وعليه أن يختار الإجابة التى يعتقد هو أنها صحيحة . مثل خارج قسمة ٧٢ ÷ ٦ هو ١٣ ، ١١ ، ١٥ ، ١٢ ، ١٤ .

٤ — فقرات المزاوجة matching items : وهنا نقـــدم للمفحوصين فقرات كل فقرة منها تتكون من عنصرين . ونضع الفقرات مصفوفة فى عمودين ، العمود الأول يشمل أحد عنصرى كل فقرة والعمود الثانى يشمل العنصر الآخر . وكل عمود له ترقيم أو ترميز (ا ب حـ د الخ) . وعلى

المفحوص أن يشير بسهم من العنصر الأول إلى العنصر الذى يتمم الفقرة ، أو أن يضع رمز العنصر المتمم أما العنصر الأول من الفقرة أمثلة :

المثال الأول : هذه عمليات ضرب والمطلوب منك أن تضع رمز الإجابة أمام السؤال :

السؤال	الإجابة
١ – ٥×٦	(ا) ٢٨
٢ – ٨×٩	(ب) ٢٧
٣ – ٧×٤	(حـ) ٤٨
٤ – ٣×٩	(د) ٣٠
٥ – ٦×٨	(هـ) ٧٢

المثال الثانى : هذه أسماء مخلوقات وأمامها أنواع الأماكن التى تعيش فيها ، والمطلوب منك أن توصل المخلوقات بالأماكن بواسطة أسهم :

تعيش الدواجن فى	كبائن
يعيش اللصوص فى	ثكنات
يعيش المصطافون فى	حظائر
يعيش الجنود فى	أوكار

المثال الثالث : هذه أسماء بلاد وأمامها أسماء عواصم والمطلوب منك وضع رقم العاصمة أمام اسم البلد الذى تقع فيه :

(ا) إنجلترا	١ – باريس
(ب) يوغوسلافيا	٢ – بودابست
(حـ) إيطاليا	٣ – براغ

(د) تشيكوسلوفاكيا	٤ ــ لندن
(هـ) رومانيا	٥ ــ بلغراد
(و) فرنسا	٦ ــ روما
(ز) المجر	٧ ــ بوخارست

مقارنة أنواع الفقرات :

لا يمكن أن نقارن بين الأنواع المختلفة من الفقرات ، التي سبق ذكرها ، مقارنة مطلقة . فلكل نوع استخدامه المختلف . كما أن لكل باحث وجهة نظر مختلفة فى تفضيل أحد الأنواع . ولكن الأستاذ Edward Green عقد مقارنة بين هذه الأنواع . وقبل أن نعرض جدول المقارنة يجب أن نعرف بأسسها . لقد اعتمد على ملاحظاته وخبرته ، وعلى المزايا والصعوبات العملية لكل نوع ، وعلى نتائج البحوث التى أجريت فى المفاضلة بين الأنواع الستة . وبناء على هذا أعطى لكل نوع تقديراً . وهذا التقدير يتراوح بين ٣ ، ١ . والرقم ١ يعبر عن أعلى رتبة بينما يعبر الرقم ٢ عن رتبة متوسطة ، والرقم ٣ يعبر عن أقل رتبة والرمز ؟ يدل على تشكك فى قيمة نوع الفقرة فى صفة معينة . وقد قارن بين الطرق الستة على أساس ١٠ مزايا . والطرق الستة هى الطرق الأربعة التى شرحناها مضافاً إليها طريقتين هما :

٥ ــ فقرات إعادة الترتيب rearangement items : وفيها تقدم للمفحوص فقرات كلماتها غير مرتبة وعليه أن يرتبها لكى تصبح جملا ذات معنى .

مثال : أعد كتابة العبارات التالية مرتبا كلماتها بحيث تصبح جملا ذات معنى :

١ ــ المطعم محمد العشاء تناول فى أمس .

٢ ــ السرير الولد على الليل ينام في .

٣ ــ الدراجة المواصلات وسيلة السيارة من أسرع من وسائل :

٦ ــ فقرات المقال Essay items : نقدم للفحوصين أسئلة يجيبون عنها في مقال ، أو نطلب مقالا بعنوان معين ، أو نعطى قضية أو مشكلة ونطلب رأى المفحوص فيها أو نسأله أن يناقشها .

مثال : ١ ــ ما رأيك في تحديد النسل .

٢ ــ اكتب مقالا في تزايد السكان .

٣ ــ تعدد الزوجات أمر موضع مناقشة ، ناقش .

ويحذر جرين من أخذ هذه المقارنة كما هي إذ أن الحكم يختلف من باحث إلى آخر ، ولكن لا بأس من سردها ما دامت مثالا يوضح رتبة كل من الأنواع الستة . على أساس نقاط عشر :

وجه المقارنة	س ، ×	التكميل	الاختبار من عدة إجابات	المزاوجة	إعادة الترتيب	المقال
سهولة الإعداد	٢	٢	٣	٢	٢	١
سهولة فهم التعليمات	١	١	١	٢	٢	١
قصر الزمن	١	٣	١	١	١	؟
صغر المسافة الخاصة بالفقرة	٢	٢	٢	١	٣	١
سهولة التصحيح ، لا عقاب جزائى	١	٢	١	١	١	٣
عدم تدخل عامل الصدفة في الإجابة	٣	١	١	١	١	٣
تعقد التفكير	؟	؟	؟	١	١	١
وضوح السؤال الفقرة ليست مشكلة	٢	٣	١	١	١	٢
اعتماد الإجابة على الاسترجاع لا التعرف	٣	٢	٣	٢	٢	١
تحليل النتائج	٣	٢	١	٢	٢	١

الجدول (١) مقارنة بين أنواع من الفقرات

يوضح هذا الجدول أن الفقرات التى يجاب عنها بصحيح ✓ أو خطأ (×) تفضل الأخرى من الفقرات فى نواح ثلاث هى : سهولة الإجراء ، وسهولة التصحيح ، وقصر الزمن اللازم للإجابة عن كل فقرة . بينما وضعت فى المرتبة الثانية من حيث : الاقتصاد فى المساحة التى تشغلها الفقرة ، وسهولة إعدادها ، ووضوحها . ووضعت فى المرتبة الثالثة من حيث : تدخل عامل الصدفة فى الإجابة ، واعتمادها على الاسترجاع أكثر من اعتمادها على التعرف ومن حيث تحليل النتائج . وتحتاج الإجابة عن فقرات صحيح ـ خطأ إلى القليل من التفكير المركب عادة ، إلا أنها أحياناً تحتاج إلى تفكير مركب ومن ثم وضعت أمامها علامة استفهام لعدم تحديد رتبتها . وفقرات الاختبار إذا وضعت بصيغة صحيح ـ خطأ تعتبر من أفضل أنواع الفقرات إذا توفرت فيها الشروط الآتية :

١ ـ عندما يكون أمام الباحث وقت قصير لإعداد الاختبار وإجرائه وتصحيحه .

٢ ـ عندما يقوم بتصحيح الاختبار كتبة مساعدون يجهلون موضوعه .

٣ ـ عندما يكون التفكير المركب واسترجاع المعلومات قليل الأهمية فى معرفة مدى (البيانات المطلوبة للإجابة عن الفقرات) .

٤ ـ عندما لا تتدخل الصدفة فى النجاح أو الفشل تدخلا خطيراً كما فى اختبارات الشخصية .

كما يوضح الجدول أن فقرات الاختبار من نوع التكميل تمتاز عن غيرها من الأنواع : بسهولة الإجراء ، وخلوها من النجاح بالصدفة وكانت رتبتها الثانية فى سهولة الإعداد، والمساحة التى تشغلها الفقرة ، وسهولة التصحيح ، واعتمادها على الاسترجاع أكثر من التعرف ، وتحليل النتائج . وكانت رتبتها الثالثة فى طول الزمن اللازم للإجابة ، ووضوح الأسئلة ، ولكن تركيب التفكير فى الإجابة عنها غير محدد . ويعتقد بعض من يفضلون استخدام فقرات

التكميل عن فقرات صحيح ــ خطأ ، أن فقرات التكميل أقل عرضة للتخمين وهى لا تشجع التفكير السطحى ؛

وفقرات الاختبار إذا جمعت بين مزايا نوعى صحيح ــ خطأ ، والتكميل تصبح من نوع الاختيار من عدة إجابات multiple type . وهى تطبق وتصحح بنفس السرعة فى تطبيق وتصحيح فقرات صحيح ــ خطأ ،وهى خالية من عامل الصدفة خلو فقرات التكميل . وفقرات الاختيار من عدة إجابات تصلح أكثر من غيرها فى اختبارات الرياضيات والهجاية والمفردات . وهى تستخدم فى الاختبارات المقننة أكثر من غيرها من الأنواع ؛

أما فقرات المزاوجــة Matching فهى تناسب الاختبارات التى تهتم بتصنيف الحقائق كما فى اختبارات الكيمياء والطبيعة . وهى توفر فرصة متوسطة لتحليل النتائج وتحاشى تدخل عامل الصدفة . وقد أعطيت أعلى رتبة فى ستة جوانب هى : صغر المساحة التى تشغلها الفقرة ، وسهولة التصحيح ، وخلوها من أثر الصدفة ، وتركيب التفكير ، ووضوح السؤال ، واعتمادها على الاسترجاع أكثر من التعرف ؛ بينما أعطيت الرتبة الثانية فى الجوانب الأربعة الأخرى .

وتتوفر فى فقرات الاختبار التى تتطلب إعادة الترتيب نفس المزايا التى تتوفر فى فقرات المزاوجة ، بل وتزيد عنها فى دقة التصحيح . ولكن التعليمات فى كل من فقرات المزاوجة وإعادة الترتيب لا يسهل فهمها سهولة فهم تعليمات بعض الأنواع الأخرى ، ولكنها عموماً تناسب تلميذ الشهادة الإعدادية فما فوق ؛

وفقرة المقال تتطلب توضيحاً شفوياً أو كتابياً أو عرضاً للحقائق . وإعدادها أسهل من إعداد أى نوع آخر من الفقرات . إلا أنها أصعب أنواع الفقرات جميعاً فى تصحيحها ، ما دام هناك تبايناً كبيراً جداً فى الإجابات

عن الفقرة الواحدة . وهى عادة تتطلب استرجاعاً وتنظيما للأفكار أكثر من غيرها . ولهذا السبب تستخدم فقرات المقال على نطاق واسع فى قياس القدرة على حل المشاكل ، وهى أيضاً مستخدمة على أوسع نطاق فى تقدير التراكيب اللغوية . وهى الطريقة الوحيدة المستخدمة فى اختبار المفحوصين الذين يحتاجون فرصة للتعبير عن أسلوبهم الفردى وخيالهم . وفى مجال الأدب لا يمكن الاستغناء عن هذه الطريقة . وما دامت المقالات الطويلة أصعب فى تصحيحها من القصيرة ، فقد اتبع تقليد فى الاختبارات المدرسية هو تحديد الإجابة فى ١٥٠ كلمة أو بمساحة معينة . ولهذا التحديد ميزة دفع المفحوص إلى مناقشة إجابته بوضوح قبل أن يكتبها .

ولا يغيب عن أذهاننا ما لنوع فقرات الاختبارات وأسئلة الامتحانات من تأثير على استعداد التلاميذ لموقف الامتحان . فإن التلميذ يوجه مذاكرته بحيث يعد نفسه للإجابة عن نوع معين من الفقرات يشمله الاختبار غالباً . كما أن المدرسين يوجهون الطلبة فى شرحهم وإرشادهم إلى الكتب الإضافية ، وفى تمارينهم يراعون أنواع فقرات الاختبار . وهذا مثال يوضح هذا الرأى نسوقه من تجربة قام بها مايرز Meyers سنة ١٩٣٦ ، وتؤيدها دراسات أخرى كثيرة : طلب الباحث من مجموعات متكافئة من التلاميذ أن تعد كل مجموعة نفسها لأحد الأنواع التالية من الامتحانات: صحيح — خطأ ، الاختيار من عدة إجابات ، التكميل ، والمقال . وعند نهاية فترة المذاكرة أعطى لكل الجماعات كل الأنواع الأربعة من الامتحانات معاً . فوجد أن التلاميذ الذين أعدوا أنفسهم لاختبارات المقال حصلوا على درجات فى كل أنواع الاختبارات متوسطها أعلى من كل المجموعات الأخرى . وحصل من أعدوا أنفسهم لاختبارات التكميل على درجات تليها ؛ بينما جاء ترتيب درجات من أعدوا أنفسهم لاختبارات صحيح

ـ خطأ ومن أعدوا أنفسهم لاختبارات الاختيار من عدة إجابات متساوياً وهو أسوأ من المجموعتين السابقتين .

وقد حاول بعض مصممى الاختبارات الجمع بين نوعين من الفقرات : فقام ماك كلاسكى Mc Clusky سنة ١٩٣٤ بالجمع بين فقرات صحيح ــ خطأ وفقرات التكميل . فطلب من التلاميذ عند وضع علامة (×) على الإجابة أن يكتبوا بجوارها الحقائق التى استندوا إليها فى الحكم بخطأ الإجابة . جعلت هذه الطريقة زمن الأجراء والتصحيح أطول مما تستغرقه اختبارات صحيح ــ خطأ ، ولكنها أعطت صورة أوضح عن معرفة التلميذ بالموضوع .

وهناك اختبارات أخرى تجمع بين طريقتى الاختيار من عدة إجابات والتكميل . فقد أعطى كيرتس Curtis ١٩٢٨ تلاميذه فقرات وأمام كل فقرة أربعة اختيارات ومساحة فارغة يستطيع التلميذ أن يكتب فيها أدق إجابة يستطيعها . هذا الجمع أطال فى زمن الأجراء والتصحيح كثيراً وخاصة فى بعض الأسئلة بالذات وبالنسبة لبعض التلاميذ بالذات . وله ميزة دفع التلاميذ إلى التفكير الأعمق ليصلوا إلى إجابة صحيحة ؛ وهذه الطريقة تفضل طريقة الاختيار من عدة إجابات . كما أنها أيضاً تمكن من مراجعة وتحسين فقرات الاختبار من نوع الاختيار من عدة إجابات ، لأن التلاميذ غالبا ما يكتبون إجابات خاطئة ، فإذا وضعت هذه الإجابات ضمن الاختيارات المرفقة بالفقرات ، فإنها تجعل فقرة الاختبار أكثر تميزا من ذى قبل . وهذا المصدر لمتغيرات الاختيارات هام لأنه يصعب عادة إيجاد ثلاثة اختيارات خاطئة تشبه الاختيار الصحيح للاجابة ، ذلك بغرض جعل تمييز الإجابة صعبا بدرجة كافية ، ذلك لأن الاختيارات الخاطئة يجب أن تكون قريبة من الإجابة الصحيحة بدرجة تجعل بعض المفحوصين يرونها مرادفة لها »

قواعد عامة فى إعداد الفقرات :

١ ـ استعمل الصيغة الإيجابية ما أمكن : فمثلا تجنب الصياغة التالية : (أنا لا أنام فى الليل كثيراً) واستخدم (أنا أنام فى الليل قليلا)

٢ ـ تجنب استعمال الكلمات التى تحمل أكثر من معنى واحد . مثل كلمة (ثقافة) وإذا استخدمتها فحدد ما تقصده بها .

٣ ـ اجعل السؤال يشتمل على فكرة واحدة مستقلة ما أمكن .

٤ ـ تجنب التعميمات المطلقة فى الزمان أو المكان لأن هذا غالبا ما يكون خطأ والمفحوص يستغل ذلك فى الإجابة . فلا تلجأ إلى عبارات مثل (دائماً) أو (فى كل مكان)

٥ ـ تجنب العبارات الشديدة الايجاز التى تجعل المفحوص شاكاً فى معناها . فلا تضع عبارة مثل (إن معانى الكلمات يتوقف على ...) ، أو (إن مشروع الثورة هو ...) بل استخدام بدلا منها عبارة مثل (أهم فائدة للسد العالى هى ...) .

٦ ـ وضح فى التعليمات الزمن المحدد للإجابة .

٧ ـ ضع تعليمات واضحة وتأكد من أنها ستتبع .

٨ ـ بالنسبة لكل الفقرات ، قم بمراجعة صعوبتها إحصائياً وادرس أثر عوامل الصدفة والغموض وسجل النتائج على بطاقة منفصلة بالنسبة لكل فقرة . هـذا سيساعدك على عمل سجل للفقرات التى تعرف صعوبتها وقيمتها :

وهناك طريقتان ناجحتان نستخدمهما فى الكشف عن الفقرات الغامضة هما : (ا) ملاحظة السلوك (ب) والتحليل الأحصائى . ولطرق الملاحظة قيمة كبيرة فى تقييم كل من اختبارات الأداء والاختبارات ا للفظية . فقد أمكن الكشف عن غموض الدرجات بملاحظ المفحوصين المختلفين أثناء أدائهم

للاختبار . ففى اختبارات التجميع Assembly واختبارات المعالجة اليدوية manipulation ، قد نجد الأفراد المختلفين يقومون بالعمل المطلوب بطرق عدة ومع ذلك يحصلون على درجات واحدة . والدرجة الموحدة هنا تمثل مهارات مختلفة ، وهكذا لا يمكن مقارنة الأفراد المختلفين على أساس درجاتهم فالدرجات هنا غامضة الدلالة عن المستوى والوظيفة .

وأفضل الطرق ، فى الحالة السابقة ، ليست الدرجات الرقمية ، ولكنها وصف طريقة قيام الأفراد المختلفين بالعمل الواحد . وكذلك لا تجدى الدرجة الدالة وحسب على الزمن المستغرق فى الأداء . ولهذا يحسن أن نحدد ونرتب موقف الاختبار أو تعليماته بحيث يقوم كل المفحوصين بخطوات واحدة بطريقة موحدة .

وفى الاختبارات اللفظية ، يسهل ملاحظة غموض التفسير . وأول طرق هذه الملاحظة ، أن يقوم الباحث بنفسه بالإجابة عن الاختبار ويسجل خبراته المحيرة . كما قد يسأل بعض الخبراء أن يخبروه كيف يفسرون كل فقرة . والتحليل الذاتى ، من هذا النوع ، عملى واقتصادى جداً . وفى تقييم الاختبارات الموضوعية التى يصممها المدرس ، يجب ، إذا أمكن ، أن يترك المدرس الاختبار أسبوعاً أو أكثر بين الإنشاء والتقييم . وهذه الفترة الزمنية تسمح للمدرس بأن يبعد جزئياً عن حالته النفسية التى تأثر بها أثناء الإنشاء أو تحيز لها والتى حددت صياغة واختيار فقرات معينة ، وهذه الفترة تمكنه من أن ينظر للأمر من زاوية أخرى . وبهذا قد يصل الباحث إلى نتائج تجعله يعدل من وجهة نظره . وبعد القيام بهذا التحليل الذاتى ، وبعد تطبيق الاختبار على الجماعة المقصودة ، يجب أن يقوم الباحث بالتحليل الإحصائى .

التحليل الإحصائى للفقرات

ثبت من الكثير من التجارب أن فقرات الاختبار التى تتطلب إجابة مختصرة ، إذا أردنا بها أن تكون فقرات جيدة ، يصعب تصميمها . وقد يكون هدف الباحث هنا أن يبتعد عن التعليمات الكثيرة المعقدة وأن ينشئ اختباراً صادقاً يستطيع الكشف عن مدى معلومات المفحوصين أو مهاراتهم . ولكن النتائج غالباً تأتى بعكس هذا . فقد وجد وود Wood سنة ١٩٢٧ أن حوالى ٣٠ ٪ من الفقرات المستخدمة فى امتحان التلاميذ للصغار كانت غير فعالة فى تمييز المستوى المنخفض من المستوى الأفضل . كما وجد Anderson أن ٣٣ ٪ من فقرات الاختبارات فى ميدان القياس التربوى كانت أكفأ من الاختبارات الكلية ، التى تشمل هذه الفقرات ، فى تقدير التحصيل أثناء الدراسة . بل لقد كانت بعض الفقرات لا تميز حتى بين الطلبة الممتازين جداً والمتخلفين جداً . والآن نناقش أسباب هذه النتائج المذهلة . على أن هم سببين لها هما (ا) استخدام فقرات غير صالحة (ب) وتدخل عوامل الصدفة .

(ا) الكشف عن الفقرات غير الصالحة :

بعد تطبيق الاختبار على جماعة التقنين ، تقيم كل فقرة بمعرفة قدرتها على التمييز بين المفحوصين تمييزاً يتفق مع تقسيمهم إلى فئات بحسب أحد أو بعض المحكات التى تقدر نفس الوظيفة التى يقيسها الاختبار . وهذا المحك قد يكون الدرجة الكلية على الاختبار نفسه أو أى درجة أو تقدير آخر نعتبره دالا . ويوضح الجدول (٢) تقدير السبعة فقرات تتطلب الإجابة عن كل منها وضع علامة ✓ أو × . والجدول عن E.B.Green .

والمقارنة فيه بطريقتين تهدفان إلى الكشف عن مدى اتفاق النتيجة أو الدرجة عن فقرة مفردة مع مقياس المحك ، والمحك هنا هو الدرجة الكلية عن الاختبار . فقسم الأفراد إلى مجموعات ثلاثة بحسب الدرجة الكلية عن الإختبار ثم حسب النسب المئوية ، فى كل قسم ، التى أجابت بصحة عن كل فقرة : والطريقة الثانية هى إيجاد معامل الارتباط بين نجاح الأفراد فى الفقرة وبين نجاحهم فى الاختبار ككل :

جدول (٢) تقييم فقرات صحيح – خطأ

النسبة المئوية الناجحة ، معاملات الارتباط

الفقرات / المفحصون مقسمون بحسب المحك	ا	ب	ح	د	هـ	و	ز
الثلث الأعلى	٪٩٥	٪٥٧	٪٩٦	٪٦٥	٪٣٣	٪٩٥	٪٨٥
الثلث المتوسط	٪٦٥	٪٥٤	٪٩٤	٪٥٧	٪٥٤	٪٩٢	٪٥٤
الثلث الأدنى	٪٥٠	٪٥٩	٪٩٧	٪٦٣	٪٦٨	٪٦٤	٪٥٦
معامل الارتباط مع الدرجة الكلية	٠ر٧٤	٠ر١٤	٠ر٠٧	-٠ر٠٦	-٠ر٣٢	ر٣٠	ر٣٢

فبإحدى الطريقتين قسم الأفراد كلهم إلى ثلاثة مجموعات : الثلث الأول هو الأعلى ، والثانى هو المتوسط والثالث هو الأدنى ، على أساس درجاتهم فى الاختبار ككل . ثم حسبت النسبة المئوية للناجحين فى كل فقرة من بين أفراد كل مجموعة من المجموعات الثلاث . والطريقة الثانية التى استخدمت ، ونتائجها مدونة فى الصف الأفقى الأخير ، هى إيجاد معاملات الارتباط بين النجاح فى كل فقرة وبين النجاح فى الاختبار ككل . فالفقرة (ا) مثلا نجح فيها

(٩٥٪) من الأفراد الذين حصلوا على أعلى الدرجات (وهم ثلث عدد المفحوصين كلهم) ، ونجح فيها (٦٠٪) من الثلث المتوسط و (٥٠٪) من الثلث الأدنى ، وكان معامل ارتباطها بالدرجة الكلية على الاختبار هو (٧٤ر) . مثل هذه الفقرة فقرة جيدة بمعنى أنها ميزت الأفراد إلى فئات تتفق كثيراً مع تميزهم إلى فئات بحسب درجاتهم على الاختبار ككل ؛ والاختبار الذى يتكون من مثل هذه الفقرة اختبار يتمتع باتساق داخلى عالى **High Internal consistancy,**

والفقرة (ب) نجح فيها أفراد الأقسام الثلاثة نجاحاً متساوياً ، فهى إذن لا تميز بين من حصلوا على درجات عالية فى الاختبار ومن حصلوا على درجات متوسطة فيه ولا من حصلوا على درجات منخفضة ، كما أن معامل ارتباطها بالنجاح فى الاختبار ككل منخفض مما يدل عل أن العلاقة عشوائية . وهذه الفقرة فى نفس الوقت ، كما يبدو من الأرقام ، صعبة على الأفراد عموماً .

أما الفقرة (حـ) فقد نجح فيها أفراد المجموعات الثلاثة بنفس النسبة تقريباً ومعامل ارتباطها قريب من الصفر . كما أنها سهلة جداً ولا تؤثر على المركز النسبي للأشخاص المفحوصين .

فى حين أن الفقرة (د) وكذا الفقرة (هـ) أعطت نتائج غامضة . فالأرقام ، فيما يتعلق بالفقرة (د) تدل على أن المجموعات الثلاثة نجحت نفس النجاح ، وإن هـــذا النجاح فوق الصدفة (٦٣٪) فهو فوق (٥٠٪) : وكان معامل ارتباطها مع الدرجة الكلية على الاختبار سالباً أى أنها لم تميز بين من حصلوا على درجات مرتفعة فى الاختبار ككل وبين من حصلوا على درجات متوسطة أو منخفضة . كذلك نجد الفقرة (هـ) غامضة الدلالة وغير مميزة ، فقد نجح فيها (٣٣٪) من أفراد أعلى مجموعة فى الاختبار بينما نجح فيها (٥٤٪) من أفراد المجموعة المتوسطة

فى الاختبار . وكان معامل ارتباطها بالدرجة على الاختبار ككل . (ــ ٣٢ ر ٠) ويمكن أن نفترض الفروض الآتية لتفسير نتائج الفقرتين (د) و (هـ) :

١ ــ أن أفراد المجموعة العليا فى الاختبار فسروا الفقرة (د) تفسيراً يختلف عن تفسير أفراد المجموعة الدنيا ، فيحتمل ألا يكونوا قد فهموا المطلوب فى الفقرة أو معناها .

٢ ــ نفس الفرض السابق محتمل فى الفقرة (هـ) .

٣ ــ أن الفقرتين (د) ، (هـ) تقيسان مهارة ترتبط ارتباطاً سالبا بالمهارة التى يقيسها الاختبار ككل .

أما الفقرتين (و) ، (ز) فقد نجحتا فى التمييز بين المفحوصين نجاحاً جزئياً . فالفقرة (و) ميزت بين الثلث الأدنى والثلث المتوسط بينما لم تميز بين الثلث المتوسط والثلث الأعلى . أما الفقرة (ز) فقد ميزت بين الثلث الأعلى والثلث المتوسط ، بينما لم تميز بين الثلث المتوسط والثلث الأدنى . ولكن الفقرتين ارتبطتا بالمحك ارتباطا موجباً ولكنه منخفض . والاختبار الذى يشمل فقرات من مثل الفقرة (و) والفقرة (ز) اختبار مقبول ، وإن كان غير مميز مثل الاختبارات التى تشمل فقرات من نوع الفقرة (ا) .

هذا النوع من التحليل الإحصائى لقدرة الفقرات على التمييز ولصعوبتها تحليل سهل وهذا سبب شيوعه . وهو يبين متى تؤثر الصدفة والغموض فى الدرجات .

النجاح بالصدفة :

١ ــ فروض نظرية .

فى الاختبارات التى تتطلب من المفحوص أن يختار إجابة من إجابتين

أو من عدة إجابات ترفق بالفقرة ، تكون هناك فرصاً لنجاح من لا يعرف الإجابة . فقد يقوم من يجهل الإجابة بوضع علامات أو إجابات مختاراً الإجابة الأولى عن كل الفقرات ، وهنا قد يحدث أن تكون الإجابة الصحيحة عن الفقرة الرابعة مثلا هى الإجابة الأولى . وبالمثل قد يقوم بعض المفحوصين باختيار الإجابات بطريقة عشوائية أو بحسب أى ترتيب دون أن يكونوا عارفين ما هى الإجابة الصحيحة .

فإذا أجاب كل المفحوصين عن كل الفقرات فى اختبار من نوع صحيح ـ خطأ أو اختبار يتكون من فقرات تكون الإجابة عنها باختيار إحدى الإجابات المدونة بعد كل فقرة ، لا يسهل تصحيح أثر الصدفة فى إنجاح البعض تصحيحاً يعدل المكانة النسبية لهذا البعض . ولهذا السبب ، يطلب بعض الفاحصين من كل المفحوصين أن يخمنوا الإجابة عن الفقرات التى لا يعرفونها . ومن الناحية الأخرى ، قد يحـــذف الأفراد الحريصون الفقرات التى يشكون فى معرفة إجابتها ، ولكن تصحيح أثر الصدفة فى النجاح عادة ما يغير مراكزهم النســبية فى الجماعة . وفى الكثير من الاختبارات المقننة ، يطلب من المفحوصين ألا يخمنوا . ولكن مهما كانت التعليمات حازمة ، يحتمل دائما أن يجيب بعض الأشخاص عن بعض الفقرات بالتخمين .

وعندما يكون على المفحوص أن يختار إحدى إجابتين ، كما فى فقرات صحيح ـ خطأ ، وتكون الفقرات الصحيحة مساوية فى عددها للفقرات الخاطئة ، يحتمل أن يخمن المفحوص متأثراً بالصدفة وحدها ، الإجابات الصحيحة عن هذه الفقرات . مختاراً الإجابة صحيح ✓ فى نصف الفقرات والإجابة خطأ (×) فى النصف الآخر . فإذا افترضنا أن الفقرات التى يخطئ فيها المفحوص ضمن الفقرات التى تشكك فى التخمين فيها ،

وبافتراض أنه تشكك فى عدد من الإجابات الصحيحة يساوى عدد الإجابات الخاطئة التى تشكك فيها ، نستطيع هنا أن نستبعد أثر النجاح بالصدفة بطرح الإجابات الخاطئة من الإجابات الصحيحة فتكون درجة الفرد هى باقى عملية الطرح .

وعندما تكون الإجابات التابعة لكل فقرة هى ثلاثة اختيارات ، يكون متوسط النجاح بالصدفة فى عدد من الفقرات هو ثلث هذا العدد ، وتكون الدرجة المصححة هى عدد الإجابات الصحيحة ناقصا نصف عدد الإجابات الخاطئة . وعندما يكون على المفحوص أن يختار إجابة من أربعة فى كل فقرة تكون درجته المصححة عن الاختبار هى عدد الإجابات الصحيحة ناقصا ربع عدد الإجابات الخاطئة . وتكون المعادلة العامة لتصحيح الدرجة عن الاختبار ككل فى حالة الإجابة بالاختيار من عدة احتمالات هى :

$$\text{الدرجة المصححة} = \text{مجموع الإجابات الصحيحة} - \frac{\text{مجموع الإجابات الخاطئة}}{\text{عدد الاختيارات فى كل فقرة} - ١}$$

$$\text{أى أن} \quad د = ج - \frac{خ}{ن - ١}$$

وهذا التصحيح ليس فعالا دائماً ، بالدرجة التى نفترضها فيه ، لأن إفتراض أن كل الإجابات الخاطئة أتت بمحض التخمين إفتراض لا يصح وخاصة فى نسبة كبيرة من الحالات . بل وأكثر من هذا ففى فقرات الاختيار من عدة إجابات لا تكون الإجابات التابعة للفقرة عادة متكافئة الصعوبة — وهو الافتراض النظرى فى تطبيق معادلة التصحيح . والشخص المحظوظ يزيد عن متوسط النجاح بالصدفة ويحصل على درجات لا يستحقها لأنه خمن الإجابة ، حتى لو طبقنا عليه معادلة التصحيح ، والشخص غير المحظوظ لا يحصل أحياناً على الدرجة التى يستحقها . ولكن الحالة الأخيرة نادرة لأن تخمين الفرد مهما كان صحيحاً لا يزيد عن الصدقة (أى لا يحصل الفرد عن أكثر من نصف الدرجة الكلية) غالباً . ولكن المفحوص عادة

ما يجد فى المعلومات والأدلة التى يقرأها فى الاختبار مرشدات توجهه فى اختبار الإجابة ، وإن كانت هذه المرشدات لا تكفى دائماً للاختيار الصحيح .

والباحث يستطيع أن يقرر ما إذا كان سيصحح أو لا يصحح أثر النجاح بالصدفة وكذا أثر الفشل بإفتراض أن لها أثرا . والباحث هنا يستند إلى الأدلة العملية التى جمعها من إجراء الاختبار وإلى الأغراض التى ينوى استخدام النتائج فى خدمتها . وإليك مثال يوضح هذا :

٢ — نتائج تصحيح أثر الصدفة :

قام رتش Ruch واستدرد Stoddoed سنة ١٩٢٥ ببحث صحا فيه نتائج تدخل الصدفة . وقد استخدما ٥ أنواع من الاختبارات التى تقيس التحصيل فى مادة التاريخ الأمريكى . وهى :

فقرات إسترجاع : بدأت الثورة الأمريكية فى سنة . . :

الاختيار من خمس إجابات : بدأت الثورة الأمريكية سنة :

١٧٦٢ ١٧٧٥ ١٧٨٣ ١٧٨٩ ١٨١٢

٣ — الاختيار من ثلاثة إجابات :

بدأت الثورة الأمريكية سنة ١٧٩٢ ، ١٧٧٥ ، ١٧٨٩ .

٤ — الاختيار من إجابتين :

بدأت الثورة الأمريكية سنة ١٧٦٢ ١٧٧٥ :

٥ — صحيح — خطأ :

بدأت الثورة الأمريكية سنة ١٧٧٥ صحيح خطأ .

طبقت صورتان من كل نوع ، صورة حوت ١٠٠ فقرة ، على مجموعات ضخمة من تلاميذ المدارس الثانوية . فكانت فقرة الاسترجاع تعطى أولا ، وتليها فقرة من أحد الأنواع الأخرى . وقد ظهر أن الإجابات

الصحيحة فى أى نوع آخر ، على نفس الفقرة ، بلغت ضعف الإجابات الصحيحة عن فقرة الاسترجاع . وعند التصحيح النظرى للنجاح بالصدفة إقترب متوسط الدرجة فى فقرة صحيح – خطأ من متوسط الدرجة فى فقرة الإسترجاع . ولكن ظهر أن النجاح بالصدفة نظريا يختلف كثيراً عن النجاح بالصدفة فعلا . كما ظهر أن الفقرات من نوع الاختيار من خمسة إجابات ارتفعت فيها درجة النجاح بالصدفة فعلا عن درجة النجاح بالصدفة نظريا . وربما ترجع هذه الزيادة إلى أن الفقرة التى حوت خمسة متغيرات قدمت معلومات إضافية هامة على هديها استطاع المفحوصون أن يختاروا الإجابة الصحيحة عنها .

كانت معاملات الثبات هى معاملات الارتباط بين نتيجة إجراء الصورة الأولى ثم الإعادة بالصورة المكافئة الثانية . وقد بلغت معاملات ثبات الدرجات التى صححت من أثر الصدفة مقادير تساوى معاملات ثبات الدرجات التى لم تصحح من أثر الصدفة .

٣ – نتائج طلب عدم التخمين من المفحوصين :

قام رتش وستودارد سنة ١٩٢٧ ببحث مشابه طلبا فيه من نصف عدد التلاميذ المفحوصين أن يخمنوا الإجابة التى لا يقطعون فيها بالتأكيد ، وطلبا من النصف الآخر ألا يخمنوا عند الشك . وعدد المفحوصين هو ٢٤٥٣ تلميذاً من أربعة فرق دراسية إعدادية وثانوية . وقد إرتفعت معاملات الارتباط بين الدرجات الخام فى أى من الصورتين المتكافئتين لدى التلاميذ الذين طلب منهم ألا يخمنوا وكان الارتفاع أكثر بقليل منه فى حالة التلاميذ الذين طلب إليهم أن يخمنوا . وهكذا تدل النتائج على أنه يحسن أن نطلب من المفحوصين ألا يخمنوا . وقد وضحت التجربة حقيقة هامة هى أن بعض المفحوصين لا يخمنوا الإجابة حتى لو طلب منهم أن يخمنوا ، كما أن بعض المفحوصين يخمنون الإجابة حتى لو طلب منهم عدم التخمين . وهذا يتضح

من حقيقة أن المفحوصين الذين خمنوا فى الصورة الأولى لم يخمنوا فى الصورة الثانية فى نفس الفقرات ولا بنفس الدرجة ولهذا كان معامل الارتباط بين الصورتين منخفضا نسبياً أى أن معامل الثبات كان منخفضا نسبياً . فلو كان هؤلاء المفحوصون قد خمنوا بنفس الدرجة وبنفس الطريقة فى نفس الفقرات فى الصورة الثانية لما اختلفت مراكزهم النسبية فى الصورة الثانية عنها فى الصورة الأولى .

وكانت معاملات الارتباط بين الدرجات فى الإختبارات التى تطلب الإختبار من عدة إجابات وبين الدرجات فى الإختبارات التى تطلبت إجابة استرجاعية متقاربة بين المجموعتين من التلاميذ (المجموعة التى طلب إليها التخمين والمجموعة التى طلب إليها عدم التخمين) . وكان الارتباط بين الدرجات المصححة وبين درجات إختبار الإسترجاع أعلى من الارتباط بين الدرجات غير المصححة ودرجات إختبار الاسترجاع . فإذا افترضنا أن الاختبارات من نوع الاسترجاع أكثر صدقا ، لانتهينا إلى أن التعليمات التى نصدرها للمفحوصين بالتخمين أو عدم التخمين لها تأثير صغير على صدق الاختبارات من نوع الاختيار من عدة إجابات .

وقد توصل توبس Toops سنة ١٩٢١ إلى نتائج مماثلة باستخدام إختبارات الحرفة Trade Tests على طلبة جامعيين ؛ كما توصل اندرو Andrew وبيرد Bird سنة ١٩٣٨ إلى نتائج مشابهة بإستخدام أسئلة من مقررات علم النفس التى درست .

ويمكن تلخيص نتائج هذه الدراسات فيما يلى :

١ ــ الأنواع الخمسة من فقرات الإختبار تختلف قليلا فى قياسها للمعلومات والمهارات المختلفة ؛

٢ ــ كل الأنواع لها معاملات ثبات مرتفعة ومتشابهة عندما تكون الفترات بين القياس الأول والقياس الثانى متماثلة ؛

٣ – التصحيح النظرى للصدفة يخفض الدرجات الكلية ، ويزيد الانحرافات المعيارية ، وهو عادة لا يغير فى الثبات أو الصدق ؛

٤ – التعليمات الصادرة إلى المفحوصين بعدم التخمين تقلل متوسط الدرجات الخام بنسبة من ١٠٪ إلى ٢٠٪ ، وتقلل متوسط الدرجات المصححة بنسبة من ٢٪ إلى ٣٪ وتزيد معامل الثبات ومعامل الصـــدق زيادة طفيفة .

٥ – الفقرات من نوع صحيح – خطأ والفقرات من نوع اختيار إجابة من إجابتين تأثرت بعامل الصدفة أكثر من غيرها ، ومن ثم زاد تأثر الدرجة الخام بمعادلة التصحيح من أثر الصدفة زيادة أكبر من غيرها ؛

والتصحيحات النظرية للصدفة لا تنطبق كثيراً على فقرات الإختبار من عدة إجابات . ولكنها تنطبق عادة على فقرات صحيح – خطأ وفقرات الإختبار من إجابتين ؛ والأسباب هى :

١ – تصحيح الصدفة يقلل الدرجات الكلية ، ومن ثم يعطى صورة أكثر دقة عن القدرة الحقيقية للفرد .

٢ – التصحيح عادة يزيد الفروق الفردية بين الدرجات الخام .

٣ – يزيد صدق الإختبار بتصحيح الدرجة من أثر الصدفة ؛

٤ – التصحيح من أثر الصدفة وقاية من النقد الموجه إلى الإختبار :

٤ – تحاشى نتائج الصدفة .

تؤدى بنا مناقشة الصدفة إلى أن نقول إنها فى الكثير من الإختبارات تكون ذات أثر صغير بحيث يمكن تجاهله ، ولكن ليست هناك طرق مؤكدة للتصحيح الكامل لأخطاء الصدفة فى الدرجات الفردية . وهناك طرق ثلاثة لتحاشى النجاح بالصدفة تستخدم فى إنشاء الفقرات والإختبارات وهى :

١ – قد يتضمن الإختبار عدداً كبيراً من الفقرات وهذا يسمح لقانون

المتوسطات Law of Averages أن يقلل أخطاء الصدفة فى الدرجة . وإذا كان الإختبار يتضمن ١٠٠ فقرة أو أكثر من نوع صحيح – خطأ ، كانت الأخطاء فى الدرجات الراجعة إلى الصدفة أصغر نسبياً عادة منها فى حالة اختبار من ١٠ فقرات من نفس النوع . فحظ الفرد أكثر عرضة للتغير خلال إختبار طويل منه خلال إختبار قصير .

٢ – يمكن استخدام الفقرات التى يختار فيها المفحوص الإجابة من أربعة أو خمسة إختيارات ، أو فقرات التكميل وفيهما يكون إحتمال الخطأ فى الإجابة كبيراً . ففقرة الإختبار من نوع الإختيار من عدة إجابات إذا تبعت بأربعة إجابات تسمح فقط بتخمين ثلث الإجابات الصحيحة وثلاثة أمثال هذا العدد إجابات خاطئة تخمينية فى المتوسط . والإختبارات المكونة من مثل هذا النوع يصغر تصحيح الصدفة فيها عادة بدرجة كبيرة بالنسبة للدرجة الكلية على الإختبار .

٣ – قد يطلب من المفحوصين أن يسجلوا ما إذا كانوا متأكدين أم متشككين فى إجاباتهم ، ثم تعطى درجات أكبر للإجابات المؤكدة . ونسبة الإجابات المؤكدة الصحيحة تدل على مدى الحرص فى موقف الاختبار . وهذا الدليل على الحرص لم يدرس فى بحوث كثيرة ، ولكن يبدو أن له دلالة عن إحدى سمات الشخصية . فقد وجد وست West سنة ١٩٢٣ ، و تروو Trow سنة ١٩٢٣ ، و جرين Greene سنة ١٩٢٩ و سنة ١٩٣٨ وجدوا أن الأشخاص يختلفون فى ميلهم إلى الحرص تبعا للعوامل الآتية :

(ا) نوع مادة الاختبار .

(ب) ألفة المفحوصين بهذه المواد .

(حـ) أمزجة Tempraments المفحوصين ودقة تدريبهم :

إذ قلت النسبة المئوية للفقرات المؤكدة الإجابة عندما كانت المادة فى الاختبار مجردة abstract بينما زادت نسبة التأكد من الإجابات الصحيحة

عن الفقرات فى حالة المواد المحسوسة Concrete . والأشخاص الذين حصلوا على درجات عالية إرتفعت النسبة المئوية لتأكدهم من الإجابة الصحيحة عن الفقرات أكثر من تأكد من حصلوا على درجات منخفضة . وقد مال التلاميذ الضعاف إلى إدعاء التأكد من الإجابة كما فعل التلاميذ الأقوياء .

أفضل طرق التصحيح

تصميم الفقرة يحدد سرعة ودقة تصحيح الإجابة . وهناك عادة عمليتان نقوم بهما قبل حساب الدرجة . العملية الأولى هى مراجعة الإجابات بمفتاح Key التصحيح والثانية هى عد الإجابات الصحيحة لحساب الدرجة الكلية . كما تعد الاجابات الخاطئة والمحذوفة Omitted عند ما يجرى تصحيح الدرجة الكلية من النجاح بالصدفة .

وفى فقرات المقال أو فقرات التكميل ، يحوى المفتاح قائمة بالإجابات الصحيحة نقارن بها الإجابات التى أتى بها المفحوص ولا يبدو أن هناك طريقة آلية مختصرة لتصجيح مثل هذه الإختبارات . ويمكن زيادة سرعة تصجيح فقرات التكميل ، بكتابة كل الإجابات التى يأتى بها المفجوص فى عامود رأسى حيث يسهل مقارنتها بالمفتاح .

أما فقرات صحيح — خطأ ، أو فقرات الإختيار من عدة إجابات ، أو إختبارات المزاوجة ، هناك ثلاثة وسائل للتصحيح : المفتاح المثقب Stencil ، والتصحيح الأتوماتيكى Automatic ، والتصـــجيح الآلى Machine Scorers وهى توفر الوقت وتزيد الدقة . والإقتصاد يزيد فى حالة البحوث التى تستخدم عدداً كبيراً من الإختبارات والتى تطبقها على عدد كبير من المفحوصين . ولا يمكن وضع قاعدة بسيطة يمكن على أساسها تحديد أكثر الطرق اقتصاداً فى التصحيح لأننا يجب أن نحافظ على

ثلاثة متغيرات على الأقل وهذه المتغيرات هى : (١) أجر الكتبة القائمين بالتصحيح (٢) تكاليف تصميم وطبع مفاتيح التصحيح الأوتوماتيكية (٣) وتكاليف تشغيل الآلات .

(ا) المفتاح المثقب :

تصحح فقرات الاختبار من من عدة إجابات وفقرات الاختبار صحيح ـ خطأ غالباً باستخدام مفاتيح مثقبة Stencil مطبوعة أو مفاتيح التصحيح التى توضع ، مطابقة تماماً ، على ورقة الإجابة أوموازية لها تماماً . والمفتاح المثقب يمكن المصحح من أن يرى الإجابات الصحيحة بسرعة ، وأن يعلم عليها بأى طريقة .

وفى الكثير من المؤسسات ، حيث يصحح عدد كبير من الإختبارات ويجرى إختبار أعداد كبيرة من الأفراد ، يطلب من المفحوص أن يجيب على ورقة إجابة منفصلة Answer sheet ثم تمر هذه الأوراق إلى ماكينة تطبع الإجابات الصحيحة بحبر ملون بجوار الإجابات التى كتبها المفحوص وهكذا يسهل مقارنة الإجابات بالمفتاح بسرعة .

(ب) التصحيح الأوتوماتيكى :

التصحيح الأوتوماتيكى يسهل عملية مراجعة الإجابات على المفتاح التى تلزمنا بها المفاتيح المثقبة . والكثير من الاختبارات طبع بطريقة تجعل وضع الإجابة يدل على مدى مطابقتها للمفتاح . وهذا المفتاح يوضع بحيث لا يستطيع أن يراه المفحوص أثناء إجراء الاختبار .

وقد نشر كلاب Clapp ويونج Young سـنة ١٩٢٩ سلسلة من الإختبارات طبع فيها المفتاح على ورقة توضع تحت ورقة الإجابة . وظهر ورقة الأجابة مغطى جزئياً بطبقة رقيقة من الكربون الناسخ . ويتكون الإختبار من فقرات يختار فيها المفحوص الإجابة من أربعة إجابات بوضع

علامة × على إحدى مربعات أربعة صغيرة مرقمة تتبع الفقرة . والمفحوص يضع علامة × بقلم عادى فى المربع الذى يعتقد هو أنه يرمز إلى الإجابة الصحيحة . وفى نفس الوقت تطبع العلامة × على ورقة المفتاح التى تبين الإجابات الصحيحة بواسطة مربعات مطبوعة صغيرة . ولحساب الدرجة على الاختبار ككل ، يقوم الباحث بعد المربعات على ورقة المفتاح التى طبعت عليها علامة × . وكان هـذا البحث أول بحث يستخدم طريقة التصحيح هذه . ولكن الطريقة الأوتوماتيكية لا تعطى الدرجة الكلية تلقائياً بل على الباحث أن يعد العلامات أو المربعات .

(حـ) التصحيح الآلى :

ماكينة العمل العالمية International Business Machine (I.B.M.) المصححة صممت سنة ١٩٣٨ وفيها تستخدم ورقة إجابة مطبوعة مصممة بدقة حيث يعلم المفحوص إجاباته بقلم خاص . وورقة الإجابة مطبوع بها خطوط صغيرة متوازية أمام رقم كل سؤال ، وهذه الخطوط تحصر بينها مسافات صغيرة ويسود المفحوص بالقلم الخاص إحدى هذه المسافات أمام كل سؤال . ولتصحيح هذه الورقة ، توضع فى الماكينة ، ثم تحركها رافعة ، وتطبع الدرجة الكلية على الورقة : والتصحيح مبنى على إتصالات كهربائية فى المواضع التى بها علامات بالقلم الخاص . وتصحيح التخمين يتم فيها وذلك بوضع مفتاح مثقب خاص فى مكان معين بالماكينة بالضغط على موضع معين مرتين . وبهذه الطريقة يمكن تصحيح ٣٠٠ فقرة صح — خطأ فى نفس الوقت . وأوراق الإجابة تجرى خلال الماكينة بسرعة وبمجرد إدخالها تصحح وتوضع الدرجة الكلية : وتحتاج هذه الماكينة إلى كاتب مدرب عليها .

الفصل السادس

الصـــدق

- معنى الصدق .
- أنواع الصدق .
- المحكات .
- حساب الصدق .
- الخطأ المعيارى للقياس .
- جمع الاختبارات فى بطارية .

الفصل السادس

الصـــــدق

معانى الصدق :

الصدق Validity مفهوم واسع له عدة معانى تختلف بحسب استخدام الاختبار ، وقد يترجم المصطلح إلى « الصحة » أو « الصلاحية » ويرجع اختلاف الترجمة غالباً إلى اختلاف المسائل التى يهتم بها الباحثون عند التعرض لمعنى الصدق . وأول معانى الصدق هو أن يقيس الاختبار ما وضع لقياسه ، أى أن الاختبار الصادق اختبار يقيس الوظيفة التى يزعم أنه يقيسها ولا يقيس شيئاً آخر بدلا منها أو بالإضافة إليها . فاختبار القدرة الميكانيكية مثلا ، لكى يكون صادقاً يجب ، بقدر الإمكان أن يقيس هذه القدرة وحدها . فلا يقيس المهارة اليدوية ، مثلا بدلا منها ولا يقيس القدرة الميكانيكية والمهارة الحركية معاً .

والاختبار الصادق اختبار يصلح لقياس الجانب المقصود أى أنه اختبار يعطى درجة تعد إنعكاساً أو تمثيلا لقدرة الفرد . فاختبار الحساب الصادق اختبار يصلح لقياس مستوى التلميذ فى الحساب . فلا يكفى أن يكون اسمه اخـــتبار حساب فالتسميات أحياناً تكون مضللة وهى أعم مما يقيسه الاختبار ، واسم الاختبار لا يدل على ما يقيسه بل على الغرض من القياس . واختبار الحساب الصادق يصلح لقياس الحساب من مستوى معين لا يجدى فى غيره كما فعل فى هذا المستوى وليكن الفرقة الدراسية المعينة . فاختبار الحساب المصمم للفرقة الثانية ، الذى ثبت صدقه بالنسبة لها ، لو طبق فى السنة الثالثة لتأثر بالذاكرة والسرعة ولو طبق فى الفرقة الأولى لبين الاستعداد

الحسابي والقدرة على القراءة والفهم . وهكذا فالاختبار يصدق بالنسبة لوظيفة معينة فى مستوى معين ، أى أنه لا وجود للصدق المطلق .

وهكذا نجد أن الصدق نسبي Relative ، بمعنى أن الاختبار يكون صادقاً بالنسبة لجماعات معينة أجريت عمليات الصدق واشتقت معاملاته من درجات هذه الجماعات أو ممثلين لها . وهنا تتحدد نسبية الصدق بخصائص وطبيعة المجتمع الأصلي الذى اشتقت منه معاملات الصدق . فالاختبار الصادق بالنسبة للأطفال ليس صادقاً بالنسبة للمسنين ، والاختبار الذى صدق على الذكور قد لا يصدق على الإناث . والاختبار الذى ثبت صدقه فى أمريكا قد لا يثبت صدقه فى بلادنا . ونقصد أن الصدق قد لا يثبت بنفس الدرجة فى هذه الأحوال جميعا . ومن أجل هذا يجب عند ذكر معاملات الصدق لاختبار معين أن نذكر طبيعة وخصائص ومستوى العينة التى استخدمت فى حسابه .

والصدق نوعى Specific ، بمعنى أن الاختبار يكون صادقا لأنه يقيس جانبا وضع لقياسه ، وهو لا يقيس جانبا آخر . فالاختبار بالنسبة لهذا الجانب الآخر غير صادق . وعند ذكر بيانات عن صدق الاختبار ، يجب أن نحدد الوظيفة أو الوظائف الممثلة فيه ، والتى يعتبر هو أداة لقياسها . وعند استخدام الاختبار نتحدد بهذه الوظائف ، فلا يستخدم اختبار H. T. P. ، مثلا ، كاختبار فنى للرسم لأنه يعتمد على الرسم كاستجابة لموقف الاختبار ، بل يستخدم فى قياس الشخصية من حيث المتغيرات التى ثبت أنه يصلح لقياسها .

وصدق الاختبار يعنى أنه يقيس ، فعلا ، الجانب الذى وضع لقياسه ، فصدق الاختبار إذن يمدنا بدليل مباشر على مدى صلاحيته لقياس أحد المتغيرات ، وهل نجح فى تحقيق الأغراض التى وضع من أجلها فى حدود الوظيفة المقاسة ، والعينة التى نقيس داخلها هذه الوظيفة . وقد عرف Kurkon الصدق بأنه « إلى أى مدى يؤدى الاختبار عمله كما يجب » : وهو

تعريف شامل يتضمن تحقيق الغرض من عملية القياس وقياس الوظيفة المعينة ، وقياسها عند نوع معين من الجماعات . وهكذا يتوقف الحديث عن صدق الاختبار على الأمور الآتية :

١ — تحديد الوظيفة أو الجانب من الحياة النفسية الذى يقيسه الاختبار .

٢ — بيانات تتعلق بالعينة التى أجريت عليها بحوث الصدق ، فيما يتعلق بما يأتى :

(ا) عدد الأفراد .

(ب) توزيعهم بحسب الجنس والسن والمهنة والتعليم و . . . الخ ؛

(جـ) تحديد خصائص وطبيعة العينة .

٣ — المعنى المحدد الذى نقصده بالصدق .

٤ — نوع الصدق .

٥ — معاملات الصدق المختلفة .

٦ — طرق حسابه .

٧ — الأخطاء المعيارية ، والأخطاء المحتملة .

٨ — تمثيل الاختبار ، باعتباره عينة من المثيرات ، للسلوك المراد قياسه :

٩ — طول الاختبار .

١٠ — معامل الثبات .

١١ — تجانس Homogeneity المجموعة .

وقد نعنى بصدق الاختبار قدرته التنبوئية ، فيعد الاختبار الصادق أداة نقيس بها عينة من المثيرات تمثل المنطقة المطلوبة من السلوك ، وتعد الدرجة دليلا نتنبأ به على نجاح من يحصل عليها عند التحاقه بتدريب أو دراسة معينة يستخدم فيها قدرة أو استعداداً قاسه الاختبار . والحق أن هذا المعنى هو الأكثر شيوعاً ، وهو معنى يسهل على أساسه حساب الصدق إحصائيا

وتجريبياً . فالصدق بهذا المعنى هو صحة توقعنا لأن ينجح فلان فى مهنة أو دراسة معينة ، لأنها تتطلب مستوى معين من القدرات والاستعدادات ، ما دام قد حصل على درجات معينة فى اختبارات تقيس هذه الاستعدادات وتلك القدرات .

وهنا يتوقف صدق الاختبار على مدى صحة التوقعات التى نقوم بها عند استخدامه ، أى على مدى مطابقة سلوك المفحوص مستقبلا لاستجاباته عن الاختبار . والمطابقة بين السلوك فى موقف الاختبار ، والإستجابة لعينة المثيرات المتضمنة فيه ، وبين السلوك فى مواقف مشابهة مستقبلا ، أمر يتوقف على مدى هذا التشابه . أى على تمثيل الاختبار تمثيلا كافياً لمجالات وجوانب وأبعاد الوظيفة التى يقيسها . فالاختبار عينة للقدرة وهو يخضع لأخطاء اختيار العينة .

فيجب ، بهذا المعنى : أن يتعلق الاختبار بالجانب المقاس وأن يتعلق الجانب بالدراسة أو المهنة التى نريد أن نتنبأ بالنجاح فيها أو الصلاحية لها . هذا التعلق relevance قد يكون كبيراً جداً كما فى حالة اختبار السائقين الذى يجرى فى قيادة سيارة فى الطريق ، فالسلوك فى موقف الاختبار هو هو السلوك فى مواقف تالية من نفس النوع . ويزيد التعلق فى اختبارات تعتبر صوراً مصغرة Job miniature للعمل . والاختبارات التى صممت بتحليل أعمال ودراسات معينة ثم أخذ عينة صادقة من متعلقات العمل أو الدراسة ووضع أسئلة تكون هذه المتعلقات إجابة عنها ، اختبارات صادقة .

والصدق هنا عملى تنبينه بالمقارنة بين درجات الأفراد فى الاختبار ونجاحهم فى العمل أو الدراسة . وقد نقوم بتحليل منطقى للوظائف والعمليات الممثلة فى الاختبار مقارنة بالعمليات والوظائف التى يعتمد عليها العمل

أو الدراسة : ويكون التحليل المنطقى للعمليات Operations قياساً كيفياً للتعلق : وهكذا يجب أن يتعلق الاختبار بالجانب المراد قياسه ، وأن يقيس هذا الجانب دائماً أى بثبات : والاتساق بين الاختبار ونفسه فى قياسه لما يتعلق به ، متمم للتعلق فى إعطاء معنى للصدق :

ومن هنا نتبين أن مفهوم الصدق يتضمن ، بمعنى ما ، مفهوم الثبات : بمعنى أن اختبار الذكاء لا يقيس الذكاء إلا إذا كان يعطى قياساً متقارباً فى المرات المختلفة ، أى أن يكون الاختبار متفقاً مع نفسه ، بينما يعنى الصدق أن يتفق الاختبار مع وظيفة يتعلق بها : ويستحيل ، بالطبع ، أن يتفق الاختبار مع ما يتعلق به أكثر مما يتفق مع نفسه : أى أنه لا يحدث أن يرتفع معامل الصدق بأكثر من معامل الثبات : وكذا فإن الاختبار الصادق اختبار ثابت فى قياساته فى المرات المختلفة ، ولكن الاختبار الثابت ، بهذا المعنى ، ليس بالضرورة اختباراً صادقاً . فإذا كان أمامنا اختبار ثابت يقيس جانباً لا نعرفه على وجه الدقة ويعطينا درجات متقاربة ، ما عرفنا أنه يقيس شيئاً ما بذاته . فالاختبار قد يزعم أنه يقيس الذكاء مثلا ويعطى للفرد الواحد عند إعادة الإجراء درجة قريبة مما حصل عليه فى الإجراء الأول ، ويعطى درجات متقاربة لو استخدمه إخصائيون مختلفون ، ولكنه برغم هذا كله قد لا يكون اختباراً صادقاً فمن أدرانا أنه يقيس الذكاء ولا يقيس الذاكرة مثلا . وقد يقيس الذاكرة فى كل مرة ويكون ثابتاً فى هذا القياس وإن قاس ما لم يرد قياسه ، أى أن الاختبار لكى يكون صادقاً يجب أن يكون ثابتاً ، وليس العكس صحيحاً ، فالاختبار قد يكون ثابتا وليس صادقاً . بل وأكثر من هذا ، فكما رأينا ، يتضمن الصدق التعلق والثبات أى أنه يكفى أن نقول إن هذا اختبار صادق فنستدل أنه ثابت ، إلى حد ما على الأقل . ومن هنا وجهت الانتقادات إلى الكثير من الاختبارات الشهيرة ذات الثبات المرتفع

فقد تبين أن كثيراً منها لا يقيس كل ما كان يزعم قياسه ، وإن قاس بعضه ، أو أنها لا تقيس هذا فقط بل تقيس شيئاً آخر إلى جواره .

وبهذه الطريقة حدد استخدام تسميات مثل اختبار الذكاء ، واختبارات الشخصية . فقد قام التحليل العاملى بتحليل اختبارات زعمت أنها تقيس كل مفهوم منهما واشتق عوامل factors نقية إلى حد ما ، وضعت اختبارات تقيسها . هذه الاختبارات تقيس أحد العوامل التى ثبت ، باستخدام منهج علمى أنه متعلق بالعمل المعين ، وبالاختبارات المتعلقة بهذا العمل ، هذه الاختبارات إذن صادقة أكثر من غيرها خاصة وأنها تقيس ، غالبا ، عوامل نقية . وكلما حدد وحلل العامل الذى صمم الاختبار لقياسه ، كان الاختبار صادقا ، وارتباط الاختبار بالعامل يحدد صدقه . وهذا هو الفهم العاملى factorial للصدق .

كما أصبح الاختبار الصادق هو الذى يرتبط ارتباطاً كبيراً ، فى درجات الأفراد عليه ، بالعمل أو الدراسة ، فى نجاح الأفراد فيها نجاحاً بدرجة معينة يطلبها واضع الاختبار أو مستخدمه ، أى أن المعنى التجريبى للصدق هو الذى يؤخذ به أكثر . وهو ، إلى جوار المعنى العاملى ، منهج علمى لدراسة الصدق بدلا من الاعتماد على ما ذكر مؤلف الاختبار من جوانب يقيسها ، وبدلا من الاكتفاء باسم الاختبار للتعرف على الوظائف التى يقيسها . فمن وجهة النظر العملية ، صدق الاختبار هو كفايته التنبؤية فى أى جانب مقاس من الحياة اليومية .

معامل الصدق :

معامل الصدق Coefficient of Validity هو أحد تطبيقات معامل الارتباط ، وهو معامل للارتباط بين درجات الأفراد فى الاختبار ودرجاتهم فى الوظائف التى يتعلق بها الاختبار ونعرف درجات الأفراد فى الوظيفة عن

ولدينا مقياس آخر غير الاختبار يقيس ما يقيسه الاختبار هو المحك criterion وسنتحدث عنه ، فيعطى الأفراد رتباً أو درجات على هذا المقياس (المحك) ونوجد بإحدى الطرق مدى ارتباط درجة الفرد فى الاختبار بدرجته فى المقياس الآخر . فإذا كان هذا الارتباط كبيراً كان معامل الصدق كبيراً وكان الاختبار صادقاً .

ومعنى هذا أن من يحصل على درجة مرتفعة فى الاختبار يحصل على درجة مرتفعة فى مقياس المحك ، وأن من يحصل على درجة منخفضة فى الاختبار يكون مستواه منخفضاً فى الوظيفة التى يقيسها الاختبار بعد أن عرفنا هذا المستوى من الدرجة على مقياس المحك . ويتراوح معامل الصدق نظرياً بين — ١ ، + ١ وإن كان لا يحدث أن ترتبط الدرجة فى الاختبار مع الدرجة فى المحك ارتباطا سالبا تاما ولا سالبا على وجه العموم . فليست هناك أحوال يحصل فيها الأفراد على درجة عالية فى الاختبار ودرجة منخفضة فى مقياس المحك دائما . أى أن معامل الصدق فعلا ينحصر بين صفر ، الذى يعبر عن عدم وجود علاقة بين الدرجة على الاختبار ودرجة على مقياس المحك تمثل مستوى الوظيفة المقاسة ، وبين واحد صحيح ، الذى يعبر عن وجود علاقة تامة موجبة بين الدرجة على الاختبار والدرجة على مقياس المحك باعتبار الأخيرة ممثلة لوظيفة يقيسها الاختبار .

والمشكلة هى أن الارتباط الموجب التام لا يحدث فى حالة الصدق ، بل نرى ارتباطاً موجباً جزئياً . فهل يكفى أى ارتباط جزئى موجب للدلالة على صدق الاختبار ؟ وما هى حدود الكسر أو النسبة المئوية لمعامل الصدق الكافى ؟ أى أن السؤال ينحصر فى تحديد مستويات يصل إليها معامل الصدق فيكون كافياً أو لا يصل فلا يكون الاختبار صادقاً . والحق أن هذا السؤال لا إجابة محددة له ولم يتفق علماء القياس على مستويات مرضية للصدق . ويمكن أن تخف حدة المشكلة وأن نحدد مستوى الصدق الذى يجب أن

يكون عليه الاختبار الذى نستخدمه ، إذا وضعنا فى الاعتبار عدة ظروف أهمها الغرض الذى سنستخدم الاختبار لأجله . وبالنسبة لمستوى صدق اختبار يستخدم فى حالات فردية تكون هناك مخاطرة فى الحكم على شخص دون تحديد مستوى ثقتنا فى هذا الحكم أى دون تحديد الخطأ المعيارى للتقدير وسنتحدث عنه عما قريب . إلا أن هناك نسبا وضعت ، على أساس دراسة إحصائية ، تبين أن استخدام الاختبار يعطى نتائج أفضل من الاعتماد على عامل الصدفة بمقدار يتوقف على قيمة معامل الصدق . فمعامل الصدق إذا بلغت قيمته ٩٠ر أعطى الاختبار نتائج تفضل الصدفة بمقدار ٥٦٪ . أى لو كنا بصدد اختبار نختار به تلاميذ يلتحقون بدراسة معينة ، ولم نعتمد على الاختبار وكانت الصدفة هى العامل الذى تحدد به الاختبار ، وجدنا ، مثلا ، ما يلى : إذا اخترنا معتمدين على الصدفة ١٠٠ تلميذ (بدون استخدام الاختبار) ، يكون عدد الناجحين ٥٠ بالصدفة أيضا ، وإذا استخدمنا اختباراً معامل صدقه ٩٠ر٠ استطعنا أن نختار ٧٨ تلميذاً ناجحين بعد استخدام هذا الاختبار .

ومعامل صدق الاختبار إذا بلغ ٨٠ر دل على أن الاختبار أدق من الصدفة بمقدار ٤٠٪ بنفس الطريقة . وإذا بلغ معامل الصدق ٧٠ر٠ كان الاختبار أفضل من الصدفة بمقدار ٢٩٪ أما إذا بلغ معامل الصدق ٤٠٪ تعطى الاختبار نتائج تفضل الصدفة بمقدار ٨٪ فقط . وفيما يلى (جدول ٢) يوضح الكفاية التنبؤية تبعاً لقيمة معامل الصدق :

معامل الصدق	الزيادة فى النسبة المئوية للكفاية التنبؤية	احتمال النجاح والفشل بين (١٠٠) يختارون بواسطة الاختبار	
		احتمال النجاح	احتمال الفشل
صفر	صفر	٥٠	٥٠
١٠ر٠	٥ر٠	٥٠ر٢٥	٤٩ر٧٥
٢٠ر٠	٢ر٠	٥١	٤٩
٣٠ر٠	٥	٥٢ر٥	٤٧ر٥
٤٠ر٠	٨	٥٤	٤٦
٥٠ر٠	١٣	٥٦ر٥	٤٣ر٥
٦٠ر٠	٢٠	٦٠	٤٠
٧٠ر٠	٢٩	٦٤ر٥	٣٥ر٥
٨٠ر٠	٤٠	٧٠	٣٠
٩٠ر٠	٥٦	٧٨	٢٢
٩٥ر٠	٦٩	٨٤ر٥	١٥ر٥
٩٨ر٠	٨٠	٩٠	١٠

(جدول ٢) الكفاية التنبؤية تبعاً لقيمة معامل الصدق

Clifford P. Froehlich and John G Darley, Studying Students — Guidance Methods of Individual Analysis, Chicago : Science Research Associates 1952

أنواع الصدق

للصدق أنواع تختلف بحسب التركيز على أحد معانيه التى سبق ذكرها . وهناك أنواع متعددة اختلفت من مشتغل بالقياس إلى آخر ، وسنقتصر على الأنواع الأكثر شيوعا واستخداما وهى :

١ — الصدق الظاهرى :

الصدق الظاهرى face Validity هو البحث عما يبدو أن الاختبار يقيسه . فبالنظر إلى الاختبار قد يبدو صادقا لأن اسمه يتعلق بالوظيفة المراد قياسها ، مع العلم بأن الاسم يدل غالبا على الغرض من استخدام الاختبار ، فالصدق الظاهرى هكذا يشير إلى كيف يبدو الاختبار مناسبا للغرض الذى وضع من أجله . ويتضح هذا النوع من الصدق بالفحص المبدئى لمحتويات الاختبار أى بالنظر إلى الفقرات ومعرفة ماذا يبدو أن تقيسه ثم مطابقة هذا الذى يبدو بالوظيفة المراد قياسها فإذا اقترب الاثنان كان الاختبار صادقا صدقا سطحيا أى بدا أن فقراته تتصل غالبا بجانب مطلوب .

وعند القيام بمثل هذا النوع من التحليل الصدقى تبدو بعض الفقرات ضعيفة أو عديمة الارتباط بالجانب المطلوب ، وهنا نستبعد هذه الفقرات . ولكن هذا الإجراء معرض لأخطاء كثيرة ، فقد يبدو اختباران ، من الناحية الظاهرية متساويين ، فى الصدق ، وعند إجرائهما يتضح الفرق . وقد حسب Guilford معامل الارتباط بين درجات مجموعة من الأفراد على اختبارين ، من الثابت أنهما صادقين ظاهريا ،وكان معامل الارتباط صفراً ، والاختباران من نوع اختبارات الترتيب بحسب الحروف الأبجدية Alphabetizing test وقد دعى هذا جيلفورد إلى أن يسمى هذا النوع من الصدق ، على سبيل التقليل من شأنه والتصغير من قيمة الاعتماد عليه ، صدق الإيمان أو العقيدة faith Validity . فمشكلة الصدق الظاهرى تتوقف على تقبل الناس لهذا الاختبار وشيوع استخدامه فى وظيفة معينة ، اعتماداً على اسمه وعلى السوابق ، وهكذا يستطيع الباحث أن يكسب ثقة الناس فى الاختبار .

أما إذا كان الاختبار ، فى نظر غير المدقق ، يبدو غير مرتبط ظاهرياً بما يراد قياسه ، اعتبر الاختبار غير مناسب للغرض الذى يراد استخدامه فيه ، ومن ثم صعب كسب ثقة الجمهور وتعاونهم فى موقف الإجراء .

وليس هذا هو الحال دائماً ، فأحياناً لا يبدو الاختبار إطلاقاً ذا صلة بالوظائف المراد قياسها ، وهنا يخفى على المفحوصين الغرض من استخدام الاختبار كما هو الحال فى الاختبارات الإسقاطية ، وبرغم هذا يكون صدق الاختبار مقبولا . أى أن الاختبار قد يصدق علمياً وعملياً ولا يصدق سطحياً وظاهرياً ، كما أننا لسنا بحاجة غالباً إلى مثل هذا النوع من الصدق ، حتى لكسب تأييد الجمهور المقاس . إلا أنه يلزم ، وخاصة عند اختبار المستويات التعليمية المتواضعة كالعمال ، أن يبدو أن نوع الفقرات والمواد المستخدمة ذات علاقة بالوظيفة المراد قياسها . بل ويحسن أن نستخدم فى الاختبار مواداً تناسب طبيعة العينة ومهن أفرادها حتى لو كنا بصدد اختبار يقيس وظيفة بعيدة عن هذه المهنة ، فيحسن فى اختبارات العلاقات المكانية إذا أجريت فى المدارس أن يطلب من التلميذ تقدير أبعاد فى الفصل أو المدرسة وهكذا . وبالمثل يفيد الصدق الظاهرى فى زيادة الدافع لموقف الاختبار .

ويفيد استخدام الصدق الظاهرى فى اختبارات التلاميذ فى الفصول ، والعمال فى الاختيار المهنى ، وفى المجال العسكرى ، وفى أغراض أخرى كثيرة : ولا يفيد فى اختبارات التشخيص أو فى المجال الإكلينيكى .

٢ ــ صدق المضمون :

صدق المضمون Content Validity هو قياس لمدى تمثيل الاختبار لنواحى الجانب المقاس . وقد يسمى بالصدق المنطقى logical Validity أو الصدق بالتعريف Validity by definition . وهنا نقوم بتحليل منطقى لمواد الاختبار وفقراته لتحديد الوظائف والجوانب والمستويات الممثلة فيه ونسبة كل منها إلى الاختبار ككل . ثم نقوم بمسح لمجال السلوك المطلوب

قياسه ، والتعرف على عوامله ، وأهمية كل عامل ، ونسبة تأثير هذا العامل على السلوك الذى يمثل الوظيفة ككل والأهمية النسبية لكل ناحية من النواحى المختلفة . ثم نطابق بين الاختبار والوظيفة التى يقيسها ، على هذه الأسس ، لنتعرف على مدى تمثيل الاختبار للوظيفة المطلوبة وعواملها ومكوناتها ونسبها . وفى كل هذا نحتاج إلى التحليل المنطقى المنظم ، والتحليل اللغوى اللفظى للأسماء ، والتعريف الدقيق للمفاهيم .

فصدق المضمون إذن يعتمد على مدى تمثيل الاختبار للمواقف أو الجوانب التى يقيسها ؛ مؤكداً على التمثيل الصادق . فإذا كان صدق مضمون الاختبار مرتفعاً كانت منطقة السلوك التى يفترض أن هذا الاختبار يقيسها ممثلة تمثيلا جيداً فى فقرات الاختبار . ويفيد صدق المحتوى فى اختبارات التحصيل والكفاية . وكما رأينا . فاختبار التحصيل الناجح يرتفع صدق مضمونه كلما كان ممثلا للموضوعات المدروسة بحسب نسبة تركيب كل موضوع منها . وسنتكلم عن الطرق المختلفة لحساب هذا النوع من الصدق .

٣ – الصدق التنبؤى :

الصدق التنبؤى Predictive Validity يقوم على أساس حساب القيمة التنبؤية للاختبار ، أى معرفة مدى صحة التنبؤات التى نبينها معتمدين على درجات الاختبار . والتنبؤ هنا يقوم على أساس أن السلوك يتمتع بدرجة من الثبات وأن الاتجاهات العامة والتنظيم فى الاستجابات يميز مجموعات من الأرجاع السلوكية لدى الشخص بميزة معينة أو بطابع خاص يظل يشكل استجاباته فى أغلب المواقف . هذا الطابع نادراً ما يتخلى الفرد عنه . وهكذا ، بقياس استجابات الفرد لمواقف معينة ، يفترض أنها تمثل قطاعاً

من المواقف التى سيعرض لها الفرد ، يمكن استنتاج كيف سيسلك المفحوص فيما بعد إذا عرفنا كيف يسلك الآن .

فإذا قسنا وظائف معينة ، ثبت أن النجاح فى مهنة أو دراسة معينة يعتمد عليها ، ثم جمعنا معلومات عن مدى نجاح من حصلوا على درجات مرتفعة ومدى نجاح من حصلوا على درجات منخفضة ، أمكننا أن نعرف إلى أى مدى سينجح الفرد الذى يحصل على درجات معينة أو ما هو مقدار احتمال نجاح من يحصل على درجات معينة . وفى التنبؤ نحتاج إلى فترة بين الإجراء وجمع البيانات عن مقياس موضوعى آخر للنجاح المقبل فى العمل أو الدراسة . أى أننا ننتظر إلى أن ينهى الطالب دراسته أو العامل تدريبه ، ثم نقارن درجاته فى الاختبار بالنجاح الذى حصله مستقبلا ونستدل على القيمة التنبؤية للدرجات المعطاه عن اختبار افترض أنه يقيس جوانب يعتمد النجاح فى المهنة أو الدراسة عليها . وسنتحدث عن طريقة حساب هذا النوع من الصدق .

٤ ــ الصدق التلازمى :

الصدق التلازمى Concurrent Validity يعنى كشف العلاقة بين الاختبار ومحك تجمع البيانات عليه وقت أو قبل إجراء الاختبار . أى أننا نقارن بين درجات الأفراد على الاختبار ودرجاتهم على مقياس موضوعى آخر يحسب مركزهم فيما يقيسه الاختبار ، على أن يعطى الأفراد درجات على المحك فى نفس الوقت الذى نطبق فيه الاختبار ، وغالبا قبل الإجراء .

٥ ــ الصدق التجريبى :

الصدق التلازمى بالإضافة إلى الصدق التنبوئى قد يشار إليهما معا بالصدق التجريبى أو العلمى أو صدق الوقائع الخارجية Emperical Validity . فهما

معا يقيسان مدى اتفاق نتائج الاختبار مع الوقائع الخارجية المتعلقة بالسلوك الفعلى فى جانب يقيسه الاختبار . والفرق بينهما ينحصر فى أننا فى الصدق التنبوئى نجمع البيانات عن المحك بعد إجراء الاختبار بفترة مستخدمين الطريقة التتبعية Follow-up method أى تتبع الأداء الواقعى للخاصية المختبرة ؛ وسنتحدث عن هذه الطريقة فيما بعد .

أما فى الصدق التلازمى فإننا نوجد مدى الارتباط بين درجات الأفراد فى الاختبار ورتبهم على مقياس المحك ، على أن يطبق المحك قبل أو أثناء إجراء الاختبار . وسنتحدث عن هذه الطريقة تفصيلا فيما بعد . وهكذا نرى أن الصدق التجريبى ، بطريقتيه التنبوئية وفيها نتتبع أفراداً قسناهم بالاختبار والتلازمية وفيها نقارن بين مجموعتين من الدرجات إحداهما على الاختبار والأخرى على مقياس المحك المطبق حين إجراء الاختبار ، يعتمد على إجراء بحوث تجريبية أو القيام بعمليات إحصائية لمقارنة مدى دلالة درجة الفرد على اختبار فى جانب ، وبين سلوكه الواقعى المعتمد على هذا الجانب

٦ — صدق المفهوم :

صدق المفهوم Construct Validity هو الارتباط بين الجوانب التى يقيسها الاختبار وبين مفهوم هـــذه الجوانب . أى أننا عند تحديد صدق المفهوم أو التكوين نقوم ، بطريقة أو بأخرى ، بتحديد ما نقصده بمصطلح يصف جانبا يقيسه الاختبار . ثم نفحص درجات الأفراد على الاختبار ونبين كيف تفسر هذه الدرجات باستخدام الجانب المقاس . فنبحث عما إذا كانت الفروق بين درجات الأفراد فى الاختبار ترجع إلى الفروق بينهم فى درجة الخاصية التى يقيسها الاختبار . فإذا كان الاختبار يقيس العصابية Neurotocism كانت الفروق بين درجات مجموعتين من الأشخاص ترجع إلى درجة قابلية كل منهما للإصابة بالعصاب (المرض النفسى) حسب تحديدنا لمفهوم العصابية .

فواضع الاختبار صاحب نظرية أو أنه صممه على أساس نظرية معينة تتعلق بجوانب افترض أن الاختبار يقيسها . وهو لكى يحدد ما إذا كان الاختبار يمثل هذه النظرية ، وما إذا كانت الفروق بين الدرجات على الاختبار تفسر بواسطة الفروق بين مستويات القدرة التى تعالجها النظرية ويقيسها الاختبار . أى أن مصمم الاختبار يحاول أن يثبت صحة النظرية التى اعتمد عليها فى إنشاء الاختبار ، وذلك بأن يحدد التنبؤات والتوقعات التى يفترضها بفحص الفروق فى الدرجات مرجعا إياها إلى النظرية التى صمم الاختبار على أساسها والتى غالبا ما تقضى بأن اختلاف الأفراد فى أحد جوانب السلوك يعنى تفاوتهم فى قدرة معينة . وهناك طريقتان لحساب صدق التكوين سنعرض لهما عما قريب .

٧ – الصدق التطابقى Congruent Validity :

هو أحد طريقتى حساب صدق المفهوم : ونحصل على معامل الصدق التطابقى بحساب مدى اتفاق درجات مجموعة من الأفراد فى الاختبار مع درجاتهم على اختبار آخر ثبت أنه صادق فى قياس نفس السمة التى يقيسها الاختبار الجديد . وسنتحدث عن هذه الطريقة فيما بعد .

٨ – الصدق العاملى :

فى الصدق العاملى Factorial Validity نهتم بحساب درجة تشبع الاختبار بالجانب المطلوب قياسه . أى أننا نبحث عن عوامل مشتركة تقيسها عدة اختبارات لنحدد مدى اشتراك هذه الاختبارات فى قياس تلك العوامل ، ومدى نقاء كل اختبار فى قياسه لأحد أو قليل من هذه العوامل . والصدق العاملى هو الطريقة الثانية لإثبات صدق المفهوم . فنجد البعض يجمع بين الصدق التطابقى والصدق العاملى عند الحديث عن صدق المفهوم . وسنتحدث تفصيلا عن طرق حساب الصدق العاملى عما قريب .

مشكلة المحكات

تعريف المحك :

المحك Criterion مقياس موضوعى مستقل ، عن الاختبار ، نقيس به صدق الاختبار . وأصل كلمة محك فى اللغة العربية يرجع إلى تقليد كان العرب يتبعونه . كانت هناك قطع من الأحجار الكريمة متفق على أنها أصيلة نقية (إذن فهى صادقة) يستخدمها الأعرابى فى فحص قطعة من الحجر ليتبين ما إذا كانت كريمة . وهو فى هذا يمسك بالقطعة الكريمة ، التى ثبت صدقها ، ويقوم بحكها (فهى محك) فى القطعة المراد اختبارها ، فإذا خدشت الأخيرة كانت غير أصيلة أى غير صادقة . ومن هنا نتبين أن المحك مقياس موضوعى ثبت صدقه نقارن به مقياسا جديدا لنتبين مدى صدقه .

الحاجة إلى المحكات :

وجدنا أننا فى حساب معامل صدق الاختبار نحتاج إلى إيجاد العلاقة بينه وبين المستوى الحقيقى للجانب الذى يقيسه الاختبار . ووجدنا أننا نصل إلى تقدير تقريبى للمستوى الحقيقى لهذا الجانب عن طريق مقياس موضوعى آخر هو المحك . ونحن إذن فى حاجة إلى المحك لمقارنة درجات الاختبار بالرتب الحقيقية للأفراد فى مجال واقعى يعتمد النجاح فيه على خاصية صمم الاختبار لقياسها . وفى الواقع نحن نعتمد على أكثر من محك واحد عند إثبات صدق الاختبار ، على أنه يجب أن تتوفر فى المحك خصائص سنعرض لها . ويجب ألا ننسى أن تواجد المحك المناسب أحياناً يكون أمراً عسيراً .

خواص المحك الجيد :

(ا) يجب أن يكون المحك متعلقاً بالوظيفة التي يقيسها الاختبار .

(ب) المحك الجيد عينة ممثلة لمنطقة السلوك المطلوب قياسها فتقدير المدرس لا يكون محكا جيداً إذا اعتمد على مجرد إعطاء تقدير للتلميذ بناءاً على موقف عارض . وكذا فتشخيص الطبيب النفسى غالباً ما يعد عينة لا تمثل كل سلوك المريض إذ أن الطبيب النفسى فى العادة أثناء التشخيص لا يعطى المريض وقتاً كافيا يلاحظ فيه سلوكه . وعلى هذا فنحن نشك فى قيمة تشخيص الطبيب النفسى كمحك .

(ح) يجب أن يكون المحك الجيد عمليا اقتصاديا . بمعنى أننا نختار محكا ملائما يسهل جمع البيانات عنه من حيث الوقت والجهد والمال .

(د) تتوفر فى المحك الجيد خاصية الثبات . أى أن المحك الجيد مقياس إذا قدرنا به فرداً واحداً عدة مرات كان التقدير متقارباً .

(ه) المحك الجيد محك موضوعى . فدرجات أو رتب الأفراد على المحك يجب أن لا تتأثر بشخص من يقوم بإعطاء الدرجة أو التقدير . فلا يجب مثلا أن نأخذ تقريراً عن عامل من صاحب عمل يقاضيه فى المحاكم .

(و) يجب أن يكون المحك مستقلا . فلا تتأثر درجة الفرد فى الاختبار بدرجته فى مقياس المحك . كما يجب ألا تتأثر رتبة الفرد فى مقياس المحك بدرجته فى الاختبار . فمن الخطأ مثلا أن نطلب من المشرف على العامل أن يقدر العامل بعد أن نحيطه علما بأن هذا العامل قد حصل على درجات مرتفعة فى اختبار الاستعداد لهذه المهنة .

أنواع المحكات :

أولا : فى الصناعة :

١ ـ كمية الإنتاج وسرعة الأداء .

٢ ـ جودة الإنتاج ودقة الأداء .

٣ ـ مقدار التلف الناشئ والفضلات المتبقية أثناء العمل .

٤ ـ عدد الحوادث والأخطاء .

٥ ـ المواظبة ودقة المواعيد .

٦ ـ الاستقرار فى العمل .

٧ ـ الأجور والمكافئات والترقيات والعلاوات .

٨ ـ المسؤولية التى عهد بها إلى العامل وعضويته فى جمعيات فنية أو نقابات مهنية .

٩ ـ مدى استفادة العامل ، أو تحصيله ، من العمل الذى يقوم به :

١٠ ـ السرعة التى ينهى بها فترة التدريب .

١١ ـ التقارير عن العمال أثناء فترة التدريب .

١٢ ـ تقديرات المشرفين والملاحظين للعامل .

١٣ ـ اختبارات الحرف .

ثانياً : فى المدرسة :

١ ـ الامتحانات النهائية .

٢ ـ الامتحانات الشهرية والدورية وامتحانات الفترات .

٣ ـ سجلات المدرسة .

٤ ـ تقديرات المشرفين ومدرسى الفصول .

٥ ـ الاختبارات التحصيلية .

٦ ــ نشاط التلميذ خارج الفصل وداخله : فى العمل ، والملعب ، والجمعيات المدرسية والهوايات والفرق الرياضية والنشاط الحر .

٧ ــ التاريخ المدرسى والتحصيلى للتلميذ .

٨ ــ البحوث والتقارير التى يقوم بها التلاميذ .

٩ ــ درجات أعمال السنة ودرجات التمارين المعطاة داخل الفصل .

١٠ ــ الاشتراك فى المناقشات والإجابة على الأسئلة وتوجيه أسئلة تدل على فهم المادة .

١١ ــ القراءات الصيفية والاشتراك فى المكتبات .

١٢ ــ المحافظة على تقاليد ونظام الجو المدرسى .

١٣ ــ المعلومات العامة .

ثالثاً فى المجالات الأخرى :

١ ــ تقديرات المفحوص ممن يعرفونه جيداً مثل الوالدين والإخوة والأقارب والأصدقاء والجيران و . . الخ .

٢ ــ البيانات الرسمية المسجلة كعدد السوابق ونوعها بالنسبة للأحداث الجانحين .

٣ ــ التقارير من المستشفيات والعيادات ومراكز التوجيه والإرشاد .

٤ ــ تشخيص الطبيب النفسى أو تقرير الأخصائى النفسى أو الاجتماعى .

٥ ــ الفروق فى العمر : ففى اختبارات الذكاء والقدرات يعتمد مفهوم العمر العقلى على أن الفرد كلما تقدم فى السن نما عقلياً .

٦ ــ مقارنة الفرد بالجماعة التى ينتمى إليها .

٧ ــ مقارنة الفرد أو الجماعة بجماعة كعينة التقنين .

٨ ــ المجموعات المتقابلة Contrasted groups : وهنا نقارن أداء أفراد في مجموعة متفوقة في أحد الجوانب المقاسة بأفراد في مجموعة تفتقر إلى هذا الجانب . كالمقارنة بين درجات العصابيين والذهانيين في مقابل درجات الأسوياء عند بحث صدق اختبارات الشخصية .

٩ ــ مقارنة الاختبار الذى نبحث صدقه باختبار آخر ثبت صدقه أجرى من قبل أو يجرى مع أو بعد الاختبار موضوع الدراسة .

١٠ ــ الإتساق الداخلي Internal Consistency أى معامل الارتباط بين نتيجة كل فقرة على حده مع نتيجة الاختبار ككل .

طرق حساب معامل الصدق

أولا الصدق الظاهرى :

بالتحليل المبدئى لفقرات الاختبار لمعرفة ما إذا كانت تتعلق بالجانب المقاس . وهذا أمر يرجع إلى ذاتية الباحث وتقديره .

ثانياً صدق المضمون :

ويقاس بالتحليل المنطقى لمحتويات الاختبار ومطابقتها مع محتويات الجانب المقاس .

ثالثاً الصدق التنبوئى :

ويقاس بإيجاد العلاقة بين الدرجات على الاختبار والدرجات على مقياس المحك الذى يطبق بعد إجراء الاختبار . وتسمى هذه بالطرية التتبعية Follow-up method . وبعد جمع البيانات عن المحك وحساب الدرجات على الاختبار نقوم بإيجاد العلاقة بينهما بإحدى طرق ثلاث هى :

(ا) طريقة النسبة المئوية Percentage method وذلك بتقسيم الأفراد على أساس رتبهم فى مقياس المحك إلى قسمين متقابلين ، كالناجحين فى مقابل

الفاشلين مثلا ، وحساب النسب المئوية من الأفراد الذين حصلوا على درجات مرتفعة أو منخفضة فى الاختبار فى كل من المجموعتين . وهكذا يكون أمامنا جدول كالآتى :

المحك / الاختبار	ناجحون	فاشلون
صفر	صفر ٪	٥ ٪
١	٣٪	١٠٪
٢	٥٪	١٠٪
٣	٤٪	٢٤٪
٤	٦٪	٢٠٪
٥	٤٠٪	١٥٪
٦	٢٠٪	٨ ٪
٧	٨٪	٥ ٪
٨	٦٪	٢ ٪
٩	٦٪	١ ٪
١٠	٤٪	صفر ٪

ومن هذا الجدول يتضح لنا أن المجموعة الناجحة وعددها ١٠٠ شخص تميل إلى أن تحصل على درجات أعلى من المجموعة الفاشلة وعددها ١٠٠ شخص ومن هذا الجدول يمكن عمل خريطة توقعات expectancy chart بمجرد النظر إلى درجة الفرد نحدد احتمال نجاحه مستقبلا ولكى نتأكد من هذا نقوم بحساب دلالة الفرق بين درجات المجموعتين من الأفراد فإذا كان الفرق جوهرياً دل هذا على أن الاختبار صادق إذ أن الاختبار نجح فى أن

يعطينا درجات بنينا على أساسها توقعات أو تنبوءات بالنجاح أو الفشل ثبت صدقها فمن حصل على درجات مرتفعة فى الاختبار سيكون غالباً ناجحاً ومن يحصل على درجات منخفضة فيه غالباً ما يفشل مستقبلا .

(ب) طريقة المتوسطات :

طريقة المتوسطات Average method تقوم على حساب دلالة الفرق بين درجات مجموعتين من الأفراد فى الاختبار ، إحداهما أخذت تقديراً مرتفعاً فى مقياس المحك والأخرى أخذت تقديراً منخفضاً فى مقياس المحك أيضاً . فإذا ثبت أن هناك فرق جوهرى بين درجات هاتين المجموعتين فى الاختبار كان الاختبار صـــادقاً ، إذ أنه يقيس سمة يعتمد العمل مستقبلا عليها .

(ج) طريقة معامل الارتباط :

طريقة معامل الاتباط Correlation method تعتمد على إيجاد معامل الارتباط بالطريقة الثنائية Biserial أو معاملاته بين درجات الأفراد على المحك ودرجاتهم على الاختبار . وهـــذه الطريقة أدق من الطريقتين السابقتين ، إذ أنها تعتمد على درجات كل الأفراد لا على متوسطهم أو نسبهم . كما أنها تمكننا من مقارنة درجات الأفراد فى أكثر من اختبار خاصة إذا أردنا أن نحدد الأهمية التنبؤية لكل اختبار نستخدمه ضمن بطارية مثلا ، غير أن طريقتى النسب المئوية والمتوسطات واضحتان فى عرض نتائجهما ومقنعتان لغير المتخصص فى الإحصاء والقياس من أمثال إدارة المدارس أو التعليم وأصحاب العمل .

رابعاً الصدق التلازمى : ويقاس هذا النوع من الصدق بإيجاد مدى ارتباط درجات الأفراد على الاختبار مع درجاتهم على مقياس المحك ، على

أن تجمع البيانات عن مقياس المحك وقت إجراء الاختبار . وتسمى هذه الطريقة بطريقة المستخدمين الحاليين The Present Employee method ، وفيها نطبق الاختبار على العمال أو التلاميذ الموجودين بالمصنع أو المدرسة ثم نجمع البيانات عن تقديراتهم على المحك وليكن كمية الإنتاج أو التحصيل ثم نحلل النتائج بطريقة النسب المئوية أو المتوسطات أو معامل الارتباط .

خامساً : الصدق التجريبي :

ويقاس بالطريقتين السابقتين معاً ، وهما الطريقة التتبعية وطريقة العمال الحاليين ، ما دام مفهوم الصدق التجريبي يتضمن مفهومى الصدق التنبؤى والصدق التلازمى . والحق أنه برغم أن حساب الصدق التجريبي يعتمد على الطريقتين معاً ، أى يحسن الجمع بينهما ، إلا أن ثمة مقارنة يمكن أن تعقد بينهما . على أن طريقتى التتبع والعمال الحاليين تستخدمان فى مجالات كثيرة وليس فى الصناعة وحدها ، ما دامت طريقة الإجراء هى هى وكذا ما دام تحليل النتائج ، بعد جمع البيانات على المحك ، تتبع الطرق الثلاث السابق ذكرها وهى طريقة النسبة المئوية وطريقة المتوسطات وطريقة الارتباطات .

فمن حيث الدافع ؛ يكون المفحوص فى الطريقة التتبعية مدفوعاً إلى جو العمل أو الدراسة الذى سيشارك فيه ، على عكس المفحوص فى حالة المستخدمين الحاليين الذى يكون قد اطمأن على وضعه فى العمل أو الدراسة والذى يعرف أن شيئاً من الاختبار لن يؤثر فى مركزه .

ومن حيث كسب ثقة وتعاون المفحوصين فى موقف الإجراء نجد أن الأفراد فى حالة التتبع يحاولون الظهور أمام المختبر بمظهر المتعاون إذ أن هذا ، فى نظرهم ، قد يؤثر على قبولهم بالمهنة أو الدراسة . بينما لا يهم المفحوص الذى يمارس العمل أو يواظب على الدراسة أن يبدو بمظهر حسن :

والطريقة التتبعية تمتاز ، فضلاً عن تقوية الدوافع وكسب ثقة وتعاون المفحوصين ، بأنها تمكن الباحث من الملاحظة الدقيقة لسلوك الأفراد فى الفترة بين إجراء الاختبار وجمع البيانات عن المحك . فالباحث هنا يستطيع أن يضع فروضاً وأن يقوم بتحقيقها على مهل وفى دقة خارجاً باستنتاجات لا يستطيعها إذا كان بصدد عمال أو طلبة يقومون بالعمل فعلا .

ولكن الطريقة التتبعية تستغرق وقتاً أطول . فقد يكون العمل أو الدراسة فى حاجة إلى الوصول إلى نتائج مباشرة وسريعة عمن يقومون بالعمل فعلا ، وهنا يتعين على الباحث أن يتبع طريقة القائمين بالعمل حالياً .

إن هناك صعوبات عملية تعترض طريقة العمال الحاليين . إذ أنه يلزم لتطبيق الاختبار أن يتوقف العمل أو الدراسة فترة تطول بحسب زمن الإجراء وقد يستحيل هذا فى بعض الصناعات . بينما نجد أن إجراء الاختبار فى الطريقة التتبعية يكون أحد خطوات الاختيار أى أنه يتم قبل الالتحاق بالعمل فعلا .

كما أن الفرد الذى يقوم بالعمل فعلا ، سواء فى الصناعة أو التعليم ، يكتسب خبرة ويحصل معلومات ويدرب قدرات وينمى مهارات تفسد اختبار بعض جوانب السلوك كالاستعدادات والميول . وهنا يكون الباحث بصدد قياس التحصيل أو الكفاية وليس الاستعداد .

وهكذا نجد أن لكل من الطريقتين عيوب ومزايا ، إلا أن الباحث يتحدد بالغرض من إجراء الاختبار عندما يختار إحدى الطريقتين وإن كان يحسن أن يجمع بينهما .

سادساً صدق المفهوم :

ويعتمد حسابه على تفسير نتائج الاختبار عن طريق النظرية التى وضع

المقياس بناءاً عليها ، أى أنه فى نهاية الأمر يرجع الفروق بين درجات الأفراد على الاختبار إلى اختلاف مستوياتهم فى الجانب الذى تعالجه النظرية ويقيسه الاختبار .

سابعاً الصدق التطابقى :

وهو أحد طريقتى حساب صدق المفهوم ويكون بحساب مدى اتفاق درجات الأفراد على الاختبار الجديد ودرجاتهم على اختبار آخر ثبت صـــدقه . ويقترح جيليكسون Gulliksen إجراء عدة خطوات عملية موضوعية لقياس صدق المفهوم بالطريقة التطابقية . يذكر ضمنها طريقة إجراء الاختبار الذى يقيس قدرة معينة ثم تنظم برامج لتدريب هـــذه القدرات لدى من أجرى عليهم الاختبار ثم إعادة إجراء الاختبار بعد انتهاء التدريب . فإذا ظهر فرق كبير ذو دلالة كان الاختبار يقيس القدرة التى دربت أى كان الاختبار يقيس ما وضع لقياسه فعلا . ومن أهم طرق حساب الصدق التطابقى ، وقد ذكرها جيليكسون أيضاً ، إجراء عدة اختبارات تقيس منطقة واحدة من السلوك ثم عقد المقارنات بين درجات الأفراد على هذه الاختبارات ومن بينها الاختبار الجديد . فإذا كان هناك معامل ارتباط (بطريقة الارتباط المتعدد Multiple correlation) كبير مع الاختبارات الأخرى الصادقة كان الاختبار الجديد صادقاً .

وتفيد طريقة إجراء الاختبار قبل التدريب Pre-training perfomance وإعادة الإجراء بعد التدريب Post-training performance فى اختبارات القدرات والتحصيل والكفاية والمهارة . ولكن هناك انتقادات عدة توجه إلى استخدام هذه الطريقة فى إثبات صدق الاختبار ما دامت درجات القدرة التى دربت قد زادت ، وأهم هذه الانتقادات هى :

هل ترجع زيادة درجات الأفراد فى الإجراء بعد التدريب إلى التدريب

وحده ؟ صحيح أن التدريب نمى القدرة أو التحصيل أو الكفاية ولكن ألم ينمو الفرد نمواً معرفياً وعقلياً ونفسياً أثناء فترة التدريب بدرجة تؤثر على أداء الاختبار مرة ثانية وتتداخل مع التدريب ؟ بمعنى أنه لو لم يعطى الأفراد تدريباً معيناً فإنهم عند إعادة الإجراء يكونون قد نموا فى مجالات متعددة فترتفع الدرجات التى يحصلون عليها برغم عدم حصولهم على تدريب رسمى formal . وهنا يكون إرجاع الزيادة الجوهرية الكبيرة فى الدرجات إلى ما أحدثه التدريب وحده أمر متعسف .

كما أن إعادة إجراء الاختبار ، فى ذاتها وبدون افتراض حدوث نمو عقلى ونفسى ومعرفى ولوجدلا ، تتأثر بعوامل التذكر والحفظ والاعتياد على موقف الاختبار ، الأمر الذى يرفع درجات الأفراد . وكأن ارتفاع درجات الأفراد ، عند إعادة الإجراء بعد التدريب أو إعطاء منهج دراسى ينمى قدرة أو جانباً يقيسه الاختبار ، لا يعود إلى فعل التدريب وحده بل وأيضاً يرجع إلى النمو development والتذكر ، وهكذا لا نتأكد أن الاختبار يقيس هذه القدرة أو الجانب وحده .

كما أن هناك مجالات من السلوك لا تجدى فيها طريقة التدريب كإثبات للصدق ، فقياس الشخصية مثلا لا يمكن أن يعتمد على هذه الطريقة . إذ أن الباحث لا يستطيع بسهولة أن ينمى « العصابية » أو « الانبساط والانطواء » مثلا .

أما عن طريقة إيجاد معامل الارتباط بين درجات الاختبار الجديد ودرجات الاختبارات التى ثبت صدقها من قبل فإنها طريقة مضيعة ، إذ أنه لا يكون هناك مبرر لإنشاء اختبار جديد ما دام سيرتبط باختبارات ثبت صدقها أى ما دام سيقيس ما قاسته هى من قبله بدقة قد يصعب عليه تحصيلها . فارتفاع معامل الارتباط هنا يعنى أن الاختبار الجديد صادق فهو يقيس ما تقيسه هذه الاختبارات ، ولكنه يعنى أيضاً أن الباحث

أحياناً لا تكون أمامه مشكلة فى قياس الجانب المطلوب بحيث يكون مدفوعاً لإنشاء اختباره خاصة وأن إنشاء الاختبارات وإثبات صدقها أمر غاية فى المشقة

ثامناً الصـــدق العاملى :

هو الطريقة الثانية ، بعد طريقة التطابق ، فى حساب صدق المفهوم ، وهو يعتمد على منهج التحليل العاملى Factor analysis . وبها نحدد مدى قياس عدة اختبارات لبعض العوامل المشتركة ، أى أننا نحدد مدى تشبع هذه الاختبارات بتلك العوامل . وتحتاج هذه الطريقة فى حسابها إلى الحصول على عدد من الاختبارات الصادقة فى قياسها لجوانب معينة ثم حساب معاملات الارتباط بين كل اختبارين من هذه الاختبارات . وهكذا يحصل الباحث على مصفوفة معاملات الارتباط Correlation matrix يمكن عن طريقها حساب درجة تشبع الاختبار بالعوامل المشتركة بين هذه المقاييس أو بين مجموعة منها . أى أن الصدق العاملى لأى اختبار هو معامل الارتباط بين هذا الاختبار وبين عامل أمكن استخلاصه بواسطة التحليل العاملى . أو هو العامل المشترك بين الاختبار وأى اختبار آخر أو مجموعة من الاختبارات تقيس نفس الوظيفة .

وتاريخ التحليل العاملى يدلنا على مدى تطور طرقه الإحصائية . وليس هذا مجال التعرض لها . ولكن تطبيق منهج التحليل العاملى فى إثبات صدق الاختبار يعتمد على طريقتين . أولاهما هى الطريقة التقليدية وتكون ، كما بينا ، بإدخال الاختبار الجديد مع مجموعة من الاختبارات . وهكذا يحصل الباحث على مصفوفة ارتباطات ، بين الاختبارات ، يحللها فيشتق عدة معاملات صدق عاملى .

ولكن لهذه الطريقة عيوب كثيرة أهمها أنه كلما كثر عدد معاملات

الصدق العاملى لاختبار واحد ، كان هذا الاختبار غير صادق إذ أنه يكون هنا مشبعاً بعدة عوامل أى أنه لا يقيس العامل الذى صمم لقياسه وحده بل يقيس إلى جواره عوامل أخرى . وهكذا يكون مضمونه موزعاً ولا يكون المقياس نقياً .

ويمكن معالجة هذا النقص بوضع تصميم يؤدى بالباحث إلى أن يكشف عن عدد من العوامل لا يزيد عن ثلث عدد الاختبارات المطبقة معا ، أى يجب أن تكون نسبة العوامل : الاختبارات = ١ : ٣ . وهكذا يتمثل كل عامل فى ثلاثة اختبارات على الأقل . كما أن من شروط تحسين هذه الطريقة أن لا يقل حجم العينة عن ٢٠٠ فرد بأى حال . وهناك مشكلة أخرى وهى تحديد مستوى الثقة فى معامل الارتباط . أى ما هو الحد الأدنى لمعامل الارتباط الذى يكفى لأن يكون جوهريا أو ذا دلالة ؟ الواقع أن هذا التساؤل غير مقطوع له بإجابة . ولكن هناك شبه اتفاق على أن معامل الارتباط هنا إذا زاد عن ٣ر٠ يكون ذا دلالة .

والطريقة الثانية فى حساب صدق الاختبار بالمفهوم العاملى تكون بإيجاد معاملات الارتباط بين فقرات الاختبار الواحد واشتقاق العوامل المشتركة بين هذه الارتباطات بين فقرات الاختبار الواحد (ويستخدم معامل الارتباط الرباعى Tetracoric Correlation) فإذا وصل الباحث إلى أن اختباره مشبع بعامل واحد فقط وأن ما دون هذا العامل بقايا ، يكون الاختبار صادقا . ولكن مفهوم الصدق هنا قريب جداً من مفهوم التجانس Homogeneity إذ أنه يصل إلى معرفة العوامل المشتركة بين الفقرات ولكنه لا يتأكد من طبيعة هذه العوامل ولا مما إذا كانت هذه العوامل هى التى قصد قياسها أم لا .

وهناك طريقة ثالثة فى حساب الصدق العاملى ، توصل إليها تيلور Taylor يقوم فيها الباحث بإيجاد معاملات الارتباط بين الدرجة على كل فقرة على حدة وبين الدرجة على بقية الفقرات معاً . وهنا يقترب مفهوم الصدق العاملى

أيضاً من مفهوم التجانس . وهذه الطريقة تفضل كثيراً طريقة حساب معاملات الارتباط بين درجات الأفراد على كل فقرة مع كل فقرة أخرى على حدة التي تتطلب حساب عدد كبير جدا من معاملات الارتباط (فإذا فرضنا أن عدد فقرات الاختبار الواحد هي ١٠٠ يكون عدد معلاملات الارتباط بالطريقة الثانية هو $\frac{١٠٠ + ٩٩}{٢}$ = ٤٩٥٠ فى حين أننا بالطريقة الثالثة نوجد ١٠٠ معامل ارتباط) . كما أن الطريقة الثانية تعتمد على إيجاد معاملات الارتباط بين كل فقرة والأخرى ، ومن المقطوع به أن معامل ثبات الفقرة الواحدة غالباً منخفض بدرجة تجعل الاعتماد على هذه الطريقة مشكوكا فيه .

وبالرغم من أننا بالصدق العاملى نستطيع الوصول إلى تنقية الاختبارات وإعداد اختبارات نقية ، إلا أن ثمة صعوبات منهجية تعترضنا . فأحيانا نوجد الارتباط بين الاختبار ومقياس محك من نوع كالتقديرات أو تشخيص الطبيب النفسى . وهنا يجب أن نحلل الاختبار تحليلا عاملياً ثم نقوم بالتحليل العاملى لمقياس المحك وبعد ذلك نطابق بين العوامل التى كشفنا عنها فى الاختبار والعوامل التى كشفنا عنها فى المحك حتى نستطيع أن نعتمد على هذا المحك . فكأننا نجرى هنا عمليات التحليل العاملى فى مرحلتين إحداهما عسيرة جداً وتخضع للذاتية فى ذلك فى حالة المحك . بينما فى حالة الاختبارات التى تستخدم كمحك نقوم بعمليات التحليل العاملى مرة واحدة وبطريقة موضوعية .

معامل الصدق معامل ارتباط

قلنا إن معامل الصدق هو تطبيق لمعامل الارتباط الذى نوجده بين الاختبار والمحك . وإليك بيان بأنواع معاملات الارتباط المستخدمة فى حساب الصدق :

(ا) معامل بيرسون Pearson : ويستخدم فى حالة المتغيرين المتصلين أى عندما تكون الدرجات على الاختبار وعلى مقياس المحك مستمرة لا انقطاع فيها continuous .

(ب) معامل الرتب Ranking : يستخدم معامل الرتب لسبيرمان إذا كان الاختبار يعطى رتباً أو درجات نحولها إلى رتب وكذا المحك .

(جـ) معامل الارتباط الثنائى Biserial : إذا كان المحك منفصلا Dicotomus مثل محك التشخيص إلى سوى وعصابى على أن تكون درجات الاختبار متصلة :

(د) معامل الارتباط الرباعى Tetrachoric : إذا كانت الدرجات على الاختبار وعلى المحك متقطعة .

(هـ) معامل الارتباط المتعدد Multiple correlation : ويستخدم فى حالة إيجاد معامل الارتباط بين درجات الاختبار ودرجات عدة محكات أخرى . كما يستخدم فى حالة تكوين البطاريات من عدد من الاختبارات .

تصحيح معامل الصدق :

قلنا إن معامل الصدق معامل ارتباط بين الاختبار والوظيفة التى يقيسها ، وأن دليلنا على مستوى هذه الوظيفة يكون غالباً مقياس المحك . وسنرى فى الفصل السابع ، عند مناقشة الثبات ، أن أى قياس يشمل ، إلى جوار قياس القدرة المقصودة ، درجة من الخطأ التجريبى فى القياس . وعلى هذا يأتى تباين ($ع^2$) أو اختلاف درجات الأفراد فى الاختبار من مصدرين ؛ أولهما هو الفروق الحقيقية بين مستويات قدرة الأفراد فى الوظيفة المقاسة . وثانيهما التباين الناشىء عن الخطأ التجريبى فى القياس . هذا الخطأ ، وبالتالى هذا التباين ، يكون فى كل اختبار أو أداة قياس على حدة .

فإذا حسبنا معامل الصـــدق بإيجاد معامل الارتباط بين

الاختبارنا والاختبار آخر يقيس نفس الوظيفة أو بينه ومقياس المحك لهذه الوظيفة ، كان لكل أداة خطأ فى القياس . وعلى هذا فنحن ندخل الخطأين معاً : وهكذا يزيد الخطأ التجريبى فى القياس ، وبالتالى فى حساب معامل الصدق ، بتداخل Overlapping الخطئين . وهذا لا يتحقق إذا كان معامل الارتباط بين كل من الاختبارين ونفسه (الثبات) كاملا أى + ١ (وهذا لا يحدث عملياً) وهكذا يتأثر معامل الصدق بمعامل الثبات ، فضلا عن تأثره بتباين الخطأ التجريبى . وعليه ، لا يكون معامل الارتباط بين الاختبار والمحك معامل صدق حقيقى true فهو يتضمن قدراً من الخطأ التجريبى . فإذا أردنا أن نحسب معامل الصدق الحقيقى أى أن تنقيته من الخطأ التجريبى ، ولذلك تسمى طريقة التصحيح للتنقية Correction for Attenuation ، نطبق المعادلة التالية :

$$ر_{ح\,س\,.\,ص} = \frac{ر_{س\,.\,ص}}{\sqrt{ر_{س\,.\,س} \quad ر_{ص\,.\,ص}}}$$

حيث ($ر_{ح\,س\,.\,ص}$) هى معامل الصدق الحقيقى للاختبار (س) أى هى معامل الارتباط بين الاختبار وبين المحك الخارجى (ص) الذى قد يكون اختباراً :

، $ر_{س\,.\,ص}$ هى معامل الصدق التجريبى الناتج من البحث وليكن (٧٠ر)

، $ر_{س\,.\,س}$ هى معامل ثبات الاختبار (س) وليكن (٩٠ر٠)

، $ر_{ص\,.\,ص}$ هى معامل ثبات المحك (ص) وليكن (٨٠ر٠)

ويكون معامل الصدق المصحح فى هذه الحالة هو :

$$\frac{٠٫٧٠}{\sqrt{٠٫٩٠ \times ٠٫٨٠}} = \frac{٠٫٧}{\sqrt{٠٫٧٢}}$$

$$= \frac{٠٫٧}{٠٫٨٤٨٧} = ٠٫٨٢$$

ويتأثر معامل صدق الإختبار بطوله والمعادلة هى :

$$r_{ط س . ص} = \frac{r_{س . ص}}{\sqrt{\frac{١ - r_{س . س}}{ط} + r_{س . س}}}$$

حيث (ط) هى طول الإختبار الجديد وليكن (١٥٠ فقرة) بالنسبة لطول الإختبار القديم وليكن ٥٠ فقرة .·. ط = ٣ ؛

، (س) هو الإختبار ؛

، (ص) هو المحك ؛

وتكون ($r_{ط س . ص}$) هى معامل الصدق بعد مضاعفة طوله ثلاثة أو أى عدد من المرات وليكن (ط) ؛

، ($r_{س . س}$) هى معامل ثبات الإختبار وليكن ٧ ر ؛

، ($r_{س . ص}$) هى معامل إرتباط الإختبار القديم مع المحك وليكن (٦٤ ر ٠) ؛

ويكون معامل صدق الإختبار بعد جعل طوله ثلاثة أمثال الطول الأصلى :

$$= \frac{\text{٠ر٦٤}}{\sqrt{\frac{\text{١} - \text{٠ر٧}}{\text{٣}} + \text{٠ر٧}}}$$

$$= \frac{\text{٠ر٦٤}}{\sqrt{\frac{\text{٠ر٣}}{\text{٣}} + \text{٠ر٧}}}$$

$$= \frac{\text{٠ر٦٤}}{\sqrt{\text{٠ر١} + \text{٠ر٧}}}$$

$$= \frac{\text{٠ر٦٤}}{\sqrt{\text{٠ر٨٠}}}$$

$$= \frac{\text{٠ر٦٤}}{\text{٠ر٨٩}} = \text{٠ر٧٢}$$

وإذا أراد الباحث أن يحدد مدى طول الإختبار الجديد بالنسبة للإختبار القديم إذا كان يريد رفع معامل صدقه إلى قيمة معينة كان عليه أن يستعمل المعادلة التالية :

$$\text{ط} = \frac{\text{ر}^{2}_{\text{ط س . ص}} \left(\text{١} - \text{ر}_{\text{س . س}}\right)}{\text{ر}^{2}_{\text{س . ص}} - \text{ر}^{2}_{\text{ط س . ص}} \times \text{ر}_{\text{س . س}}}$$

حيث ط طول الإختبار الجديد بالنسبة للإختبار الأصلى (س)

، $ر_{\text{ط س ص}}$ معامل الصدق المطلوب وأن $ر_{\text{س . ص}}$ هى معامل صدق الإختبار الأصلى وأن $ر_{\text{س . ص}}$ هى معامل ثبات الإختبار الأصلى وهى نفس الرموز المستعملة فى المعادلة السابقة .

وبالتعويض فى هذه المعادلة بالمقادير المستعملة فى المثال السابق نجد أن

$$\text{ط} = \frac{(\text{٠ر٧٢})^{٢}\,(\text{١} - \text{٠ر٧})}{(\text{٠ر٦٤})^{٢} - (\text{٠ر٧٢})^{٢}\,(\text{٠ر٧})}$$

$$= \frac{(\text{٠ر٥٢})\,(\text{٠ر٣})}{(\text{٠ر٤١}) - (\text{٠ر٥٢})\,(\text{٠ر٧})}$$

$$= \frac{\text{ر١٥٦}}{\text{ر٠٥}}$$

$$= \text{٣ تقريباً}$$

أى أن طول الإختبار المطلوب هو ٣ أمثال الإختبار الأصلى .

إستخدام الصدق فى حالات أخرى

(١) التنبؤ بالحالات الفردية :

إذا قلنا إن معامل صدق الإختبار هو ٧ر٠ وكنا نعنى الصدق التنبؤى وكان هذا المعامل هو معامل الإرتباط بين الإختبار الجديد وإختبار صادق قديم فما معنى هذا ؟ إلى جوار كل المعانى التى سبق التعرض لها يدلنا هذا الرقم على أننا فى ٧٠ ٪ من الحالات نصدر حكماً تنبؤياً سليما باستخدام هذا

الإختبار . أى أنه لو كنا بصدد اختيار تلاميذ وتقدم لنا ١٢٠٠ تلميذ واخترنا باستخدام الإختبار ١٠٠٠ تلميذ كان تنبؤنا بنجاح ٧٠٠ تلميذ منهم تنبؤاً صحيحاً . ولكننا إذا كنا بصدد حالات فردية كان الأمر أشق وكانت هناك مخاطرة فى الحكم على الفرد بهذا الإختبار أو بأى إختبار آخر ، إذ أننا نكون معرضين للخطأ بإحتمال ٣٠ ٪ . هذا فضلا عن أنه أثناء عمليات القياس المختلفة نكون قد تعرضنا لخطأ نسميه الخطأ المعيارى .

فإذا عرفنا هذا الخطأ المعيارى للدرجة التى أعطيناها للفرد أمكننا أن نضع حداً أعلى لا تزيد درجة الفرد الحقيقية عنه وحداً أدنى لا تقل عنه : هذا فى أغلب الأحوال خاصة وأن معامل الارتباط ، الذى نحسبه فى إيجاد أى معامل صدق ، يحسب من درجات عينة التقنين وليس من درجة فرد واحد مما يجعل التنبؤ بدرجة الفرد عرضة للخطأ . فإذا فرضنا أن :

الخطأ المعيارى لدرجات المحك هو $ع_{ص}$

الخطأ المعيارى لدرجة الفرد هو $ع_{ت}$

معامل الإرتباط بين الإختبار والمحك هو $ر_{س ص}$

باعتبار أن س ترمز إلى الإختبار ، ص ترمز إلى المحك .

إذن $ر_{س ص}$ هى معامل الصدق .

وهكذا يكون الخطأ فى تقدير الدرجة ورمزه $ع_{ت}$ مساوياً للمقدار الآتى

$$ع_{ت} = ع_{ص} \sqrt{١ - ر^{٢}_{س ص}}$$

مثال : فى الإختبار س حصل أحد المفحوصين على الدرجة ١٠ ووجدنا أن الخطأ المعيارى لدرجات المحك هو ٠,٢٥ وكان معامل الإرتباط بين المحك

والإختبار (أى معامل صدق الإختبار) هو ٠,٨ فتكون درجة الفرد الحقيقية فى هذا الإختبار مستبعدين بذلك كل أخطاء القياس والمقياس محصورة بين ما يلى :

$$\text{ع}_{\text{ت}} = \text{ع}_{\text{ص}} \sqrt{١ - \text{ر}^{٢}_{\text{س ص}}}$$

$$= ٢٥ر٠ \sqrt{١ - (٨ر٠)^{٢}}$$

$$= ٢٥ر٠ \sqrt{١ - ٦٤ر٠}$$

$$= ٢٥ر٠ \sqrt{٣٦ر٠}$$

$$= ٢٥ر٠ \times ٦ر٠$$

= ١٥ر٠ من الدرجة

وحيث أن درجة الفرد هى ١٠ فيكون الخطأ فى تقديرها = ٠,١٥ × ١٠ = ٥ر١

إذن درجة الفرد الحقيقية تنحصر بين ١٠ + ٥ر١ و ١٠ — ٥ر١ أى بين ١١,٥ ، ٨,٥

(ب) فى التصفية :

فى التصفية تستخدم جداول تيلور ورسل Taylor and Russel التى تحدد النسبة المئوية التى يمكن إختيارها من بين المتقدمين على أساس صدق الإختبار ونسبة الناجحين من الأفراد الذين اخترناهم دون الاعتماد على الإختبار .

(حـ) فى تكوين البطاريات :

تتكون البطاريات من عدد من الإختبارات التى تقيس عدة جوانب تتعلق بعمل أو مهنة أو دراسة أو . الخ . ولكن يجب أن تقيس هذه

الإختبارات عدة جوانب لا تختلف حتى تتعارض ولا تقترب حتى تتطابق . ويحصل الفرد على درجة كلية على البطارية أو على درجات فى كل إختبار منها . ويجب أن تحدد أوزان Weights أى مقادير بمعادلات للتصحيح لكل إختبار حسب أهمية الجانب الذى يقيسه بالنسبة لمجموع الجوانب التى تقيسها البطارية . وهكذا تكون للفرد الواحد عدة درجات تمثل بيانياً فى صفحة نفسية Psychoprofile تعبر عن صورة كلية لأحد مجالات السلوك أو الحياة النفسية .

الفصل السابع

الثبـــات

- مقـــدمة .
- إعتبارات منطقية فى تقدير الثبات .
- الثبات وتحليل التباين .
- موضوعية القياس .
- عمليات حساب الثبات .
- تفسير معامل الثبات .
- العوامل المؤثرة على ثبات الإختبار ..
- معامل الثبات ومعامل الصدق ..
- بيانات الثبات .

الفصل السابع

الثبـــات

مقدمة :

يقوم التعريف بأى إختبار على ذكر معلمين أساسيين فيـــه هما ثباته وصدقه . وكلمة « إختبار ثابت » كترجمة للمصطلح « Reliable » ليست كافية تماماً . فالأصل الأجنبى للكلمة يدل على أن هذا الإختبار « موثوق فيه ويعتمد عليه » وهذا هو المقصود غالباً بكلمة ثابت هنا . وقد يكون هذا المعنى هو النتيجة التى تصل إليها باستخدام إختبار يعطى تقديرات ثابتة . أى أنه فى عملية القياس لو كررنا الإجراء لحصلنا على نتائج متسقة Consistant عن الفرد ، بمعنى أن درجته لا تتغير جوهرياً بتكرار إجراء المقياس عليه ، أو بمعنى أن موقف الفرد بالنسبة لجماعته لا يتغير جوهرياً بين مرات الإجراء المتعددة . وهنا يقترب معنى الثبات من مفهوم الإتساق ، إتساق نتائج إجراء الإختبار مع نفسها بين مرات الإجراء المتعددة .

والثبات Reliability قد يعنى الإستقرار Stability ، بمعنى أنه لو كررت عمليات قياس الفرد الواحد لأظهرت درجته شيئاً من الإستقرار . كما أن الثبات قد يعنى الموضوعية objectivity بمعنى أن الفرد يحصل على نفس الدرجة كائناً من كان الأخصائى الذى يطبق عليه الإختبار أو الذى يصححه . وهنا يكون الإختبار الثابت مقياساً يقدر الفرد تقديراً لا يختلف فى حسابه إثنان ، تماماً كالترمومتر الذى لا يختلف فى قراءته إثنان .

ومعامل الثبات هو معامل الإرتباط بين درجات الأفراد على الإختبار بين مرات الإجراء المختلفة ، أو بين تقديرات من يقومون بتصحيح الإختبار فى المرات المختلفة ، أو بين نتائج إجراء الإختبار على مجموعة واحدة من

الأفراد على أن يقوم بالإجراء إخصائيون مختلفون . أى أن معامل الثبات هو معامل الإرتباط بين الإختبار ونفسه . فنحن نحسب معامل الثبات من معامل الإرتباط بين الدرجات على الإختبار فى مرات الإجراء المختلفة سواء كنا نعيد الإجراء بنفس الإختبار بنفس الصورة Form أو بصورة مكافئة لها Equal form . ويتضح قولنا أن معامل الثبات هو معامل الإرتباط بين الإختبار ونفسه فى أحد طرق حساب الثبات ، ونقصد عند حساب الإرتباط بين درجتى الفرد الواحد على نصفى الإختبار . وسيتضح هذا عند مناقشة طرق حساب معامل الثبات .

ومعامل الثبات ، بما أنه معامل إرتباط ، لا يكون كاملا أو لا شىء it is not all or none . أى أن معامل الثبات نسبى relative فهو يختلف تبعاً لمتغيرات عدة سنذكرها فيما بعد . كما أن معامل الثبات لا يصل إلى ١ صحيح كما لا يصل إلى صفر . والمشكلة هى نفس المشكلة التى صادفتنا فى معامل الصدق : متى يكون معامل الثبات كافياً للإعتماد على نتائج الإختبار ؟ ليست هناك قاعدة تحدد قيمة معامل الثبات الكافى . ولكن لا شك فى أن قيمة معامل الثبات مع غيرها من خصائص الإختبار وأهمها الصدق ، هى التى تحدد درجة دقة الحكم ومدى الاعتماد على الإختبار . ومن هنا نتبين أهمية معرفة قيمة معامل ثبات الإختبار الذى نستخدمه .

وأيا كان الجانب الذى نقيسه ، علما أو لغة أو حساباً ، نكون معرضين لتدخل عامل الصدفة . ويحوى الإختبار خطأ الصدفة Chance error بدرجات متفاوتة ، ولكنه لا يخلو من خطأ الصدفة هذا أبداً . ويختلف خطأ الصدفة بإختلاف الفاحص والمفحوص وموضوع القياس . وإذا كان خطأ الصدفة صغيراً زاد معامل ثبات المقياس . وإذا زادت نسبة أخطاء الصدفة بإختلاف الأفراد قل ثبات المقياس .

قلنا إن الثبات نسبى وكنا نقصد مدى إختلاف المفحوصين عن أنفسهم بين مرات الإجراء المختلفة . فنحن نحسب معامل الثبات بناءاً على إرتباط

درجات مجموعة من الأفراد بنفسها أى درجاتهم فى المرة الأولى مقارنة بدرجاتهم فى المرة الثانية بنفس الصورة أو بصورة مكافئة من الإختبار . وهنا نكون معرضين لخطأ . هذا الخطأ هو أن هؤلاء الأفراد أنفسهم يتغيرون . وعلينا أن نعرف مدى تغيرهم حتى نعرف سبب إختلاف درجاتهم فى المرات المختلفة فقد لا يكون التغير فى الإختبار بل فيمن يطبق عليهم الإختبار .

وهناك ثبات مطلق Absolute . وتحدد درجة هذا الثبات المطلق بمدى التغير الحقيقى فى نتائج إجراء الإختبار على نفس الفرد . وتسجل درجات الفرد فى كل مرة ثم يحسب الإنحراف المعيارى لها ويسمى هذا الإنحراف بالخطأ المعيارى للقياس Standard Error of Measurement .

إعتبارات منطقية فى تقدير الثبات

يتضمن تقدير ثبات أداة القياس نوعين من العمليات ، عمليات تجريبية وأخرى إحصائية . فمن ناحية يلزمنا أن نجرى المقياس على مجموعة محددة من الأفراد متبعين خطة تجريبية محددة ومحافظين على شروط التجريب المحددة . ومن ناحية أخرى ، تحال درجات الأفراد فى هـذا الإجراء بإستخدام أنسب الطرق الإحصائية التى تلائم مميزات ثبات الإختبار .

وتتضمن الخطوات التجريبية فى تقدير الثبات جوانب منطقية تتعلق بمشكلة القياس ، بحيث أنه يجب علينا أن نحلل الأغراض التى نحسب ثبات المقياس من أجلها ، فيجب أن نخطط العمليات التجريبية وأن تقيم من خلال هذه الأغراض :

الثبات وتحليل التباين:

عند إجراء أداة القياس على مجموعة من الأفراد وإعطاء أفرادها درجات ، نجد أن توزيع هذه الدرجات يشمل مدى كبيراً وسبب وجود

هذا المدى من الإختلاف ، أو إنحراف الأفراد عن متوسطهم أى التباين Variance ، يرجع إلى عدة عوامل متدخلة . وبعض هذه العوامل يرجع إلى وجود فروق حقيقية بين أفراد الجماعة فيما يتعلق بالصفة التى نقيسها . والبعض الآخر من العوامل يشير إلى عدم الدقة فى قياس الأفراد كل على حدة . وتقدير ثبات أى مقياس يعتمد على تحديد مدى إختلاف درجات الأفراد للفروق الحقيقية بينهم ، ومدى دقة قياس الأفراد المعينين .

وقد إشتق الإحصائيون معاملات تفيد فى وصف مدى تغير أو إختلاف مجموعة من الدرجات . وأفضل هـذه المعاملات هو معامل التباين ويرمز إليه بالرمز ($ع^٢$) . ويعرف بأنه متوسط مربعات انحرافات درجات جماعة من الأفراد عن متوسط هذه الجماعة وهو يساوى $(\frac{مح^٢}{ن})$. حيث ح هى الفرق بين درجة الفرد ومتوسط الجماعة ، ن هى عدد أفراد الجماعة .

ومعامل التباين له ميزة إمكان جمع تباين الجماعات الصغيرة للحصول على تباين الجماعة الكلية وكذلك طرح تباين جماعه جزئية من الجماعة الكلية للحصول على تباين الجماعة الكلية الجديدة بعد حذف هذه الجماعة الجزئية منها ، كما يمكن جمع تباين نفس المجموعة من الأفراد بالنسبة لخاصية معينة على تباينهم بالنسبة لخواص أخرى . فمثلا يمكن حساب التباين فى المجموع الكلى للدرجات لفرقة من التلاميذ بجمع تباينهم فى مواد مختلفة مثل الحساب والعلوم واللغة . . الخ فيكون التباين الكلى هو :

$$ع^٢ = ع_١^٢ + ع_٢^٢ + ع_٣^٢ +$$

حيث $ع_١^٢$ هى تباين المجموعة بالنسبة للحساب ، $ع_٢^٢$ تباينهم بالنسبة للعلوم ، $ع_٣^٢$ تباينهم بالنسبة للغة . . . الخ وحيث $ع^٢$ هى تباينهم بالنسبة للمجموع الكلى لهذه المواد .

وفى حالة القياس فى العلوم الطبيعية يمكن إعادة القياس والحصول على

نفس الدرجات وحساب التباين فتظل قيمته كما هى بالنسبة لنفس الخاصية . أى أنه يمكن إستعادة Reproducibility نفس التباين : أما بالنسبة للموضوعات النفسية والتربوية ، فإن النتائج تختلف بإختلاف المرات . ويسمى الخطأ أو الفرق بين قيمة التباين فى المرات المختلفة بخطأ التباين Error variance . هذا الخطأ يشير إلى إنخفاض ثبات المقياس .

فالقيمة العددية لمعامل ثبات الإختبار تتوقف على تباين الدرجات على تباين الدرجات فى هذا الإختبار على أن يرجع هذا التباين إلى الفروق الحقيقية بين الأفراد فى الصفة التى يقيسها الإختبار . ولا يكون الإختبار ثابتا عندما يكون خطأ التباين فيه كبيراً . ونعود فنقول إن المشكلة الأساسية فى تحديد ثبات عمليات القياس هى مشكلة تحديد التباين الحقيقى بين الأفراد وتحديد مدى خطأ التباين الذى نقبله .

أسباب تباين درجات الأفراد فى الإختبار :

يهمنا الآن أن نعرف لماذا تتباين أو تختلف درجات الأفراد فى الإختبار ، وهل يرجع التباين إلى الإختبار نفسه أم إلى المفحوصين أم إلى القائم بالإجراء أو التصحيح . فإذا تعرفنا على مصادر التباين أمكننا أن نعالجها فنزيد ثبات الإختبار . وتختلف درجات الأفراد ، عن أدائهم للإختبارات ، بحسب عوامل ذات طبيعة مختلفة أتت من مصادر متعددة : وليست القائمة التالية حصرا لكل هذه العوامل وإنما تمثيلا لأهمها : فالتباين فى درجات الأفراد قد يرجع إلى أحد أو بعض أو كثير من العوامل الآتية :

١ – مستوى الفرد فى سمة عامة ثابتة تؤثر فى أدائه لعدد من الإختبارات ، مثل القدرة اللفظية :

٢ – درجة قدرة الفرد على الإجابة تتأثر بطريقته فى فهم التعليمات.

وطريقة الأداء ومعالجة المواد . وقد تكون للفرد دراية خاصة بنوع من الخبرات والمهارات التى تعينه على الإجابة .

٣ – قد يكون الفرد ملما ببعض الحقائق أو المعارف الصعبة وغير ملم بالأسهل منها . وقد يحدث أن يطلب الاختبار منه ذكر ما حدث مصادفة إن كان يعرفه ؛

٤ – مستوى دوافع الفرد للإجابة عن الاختبار .

٥ – الحالة الجسمية والعقلية والوجدانية للفرد وقت الإجراء .

٦ – قدرة المفحوص على التزييف والغش والخداع .

٧ – الظروف الخارجية أثناء الإجراء كالضوضاء والتهوية والحرارة .

٨ – ألفة الفرد بالإجابة عن نوع معين من الاختبارات أو بمعالجة الأجهزة والمواد المستخدمة فيها .

٩ – اضطراب الذاكرة أو الانتباه .

١٠ – تدخل عامل الصدفة عندما يختار الفرد بين عدة إجابات فقد يخمن الإجابة الصحيحة فيعطى عنها درجة فى مرة ثم يخمن إجابة مخالفة لنفس الفقرة ، فى إجراء تالى ، فلا يحصل على نفس الدرجة .

١١ – بعض السمات غير ثابتة بطبيعتها ، وخاصة سمات الشخصية .

وكل هذه المصادر ترجع إلى المفحوص نفسه : ولكن التباين قد يرجع إلى أسباب أخرى ، مثل الاختبار نفسه ، سنعرض لها فيما بعد .

موضوعية القياس :

لكى نوفر الدقة فى تقديراتنا ، يجب أن نفترض أن كل درجة يحصل عليها المفحوص تعد تمثيلا لقدرته الحقيقية متأثرا ببعض الأخطاء التجريبية فى عمليات القياس فتكون :

الدرجة على الاختبار = الدرجة الحقيقية ± الدرجة الناتجة عن الخطأ التجريبي فى القياس .

أى أن $س_ق = س_{\infty} \pm س_خ$

وبالمثل فإن تباين الدرجات على الاختبار = تباين الدرجات الحقيقية ± تباين الدرجات الناتجة عن الخطأ التجريبى فى القياس .

أى أن $ع^٢_ق = (ع_{\infty} \pm ع_خ)^٢$

فإذا كان الخطأ التجريبي فى القياس صغيرا جداً ، قاربت الدرجات على الاختبار الدرجات التى يجب أن يحصل عليها الأفراد الممثلة لمستواهم الحقيقى فى القدرة المقاسة) . وهنا يرجع اختلاف درجات الأفراد فى الإختبار ($ع^٢_ق$) وحسب إلى الفروق الفردية الحقيقة بينهم فيما يتعلق بالجانب المقاس ($ع^٢_{\infty}$) ويصبح التباين فى درجات الإختبار مساويا للتباين فى مستويات القدرة .

أى أن $ع^٢_ق \doteq ع^٢_{\infty}$

ولكننا عادة لا نعرف مقدار المستوى الحقيقى لقدرات الأفراد . ولذا فنحن نطرح تباين الدرجات الناتجة عن الخطأ التجريبي فى القياس من تباين الدرجات فى الاختبار . وبالمثل يصبح التباين الحقيقى هو باقى طرح التباين بين الدرجات الناتجة عن الخطأ التجريبى فى القياس من التباين بين درجات الأفراد فى الإختبار .

وبما أننا عند إعطاء الفقرات درجات على الإختبار سلمنا بأن ثمة خطأ تجريبي يتدخل ، وبما أنه يهمنا أن نعرف ما هى نسبة الخطأ التجريبى فى الدرجة على الإختبار ، فانه يلزمنا أن نعرف تباين الدرجات الناتجة عن الخطأ التجريبي فى القياس حتى نعرف مدى الاختلاف أو التباين الحقيقى

(الذى يعبر عن فروق فردية حقيقية بين مستويات قدرة الأفراد) .
وهكذا نستطيع أن نحدد نسبة التباين الحقيقى المتضمن داخل تباين الدرجات المعطاة على المقياس ، وهذا يوصلنا إلى تحديد قيمة معامل الثبات . أى أن معامل الثبات هو ، بهذا المعنى ، جزء من التباين الحقيقى الداخل فى تباين الدرجات التجريبية ، وهو يساوى :

$$ر_{ق ق} = \frac{ع^2_{\infty}}{ع^2_{ق}}$$

حيث $ر_{ق ق}$ هى معامل الثبات أى معامل الارتباط بين الدرجات على الاختبار (وهى ق) فى مرتى الإجراء أو صورتى الاختبار أو نصفى المقياس .

وحيث $ع^2_{\infty}$ هى التباين الحقيقى وهو مربع الانحراف المعيـــارى لدرجات الأفراد التى تمثل قدرتهم فعلا . (وهى درجة افتراضية قد نصل إليها عن طريق إجراء مقياس كامل الدقة ، أو يفترض أنها متوسط درجات الفرد الواحد لو طبق عليه الاختبار عددا لا نهائيا من المرات ، أو بتطبيق عدد كبير جداً من الاختبارات التى تقيس وظيفة معينة على نفس الفرد) . أى أنها التباين الحقيقى .

وحيث $ع^2_{ق}$ هى مربع الانحراف المعيارى لدرجات الأفراد على الاختبار أى أنها تباين درجات المقياس .

وهذا مثال يوضح طريقة الحساب :

إذا فرضنا أننا طبقنا الاختبار مرتين أو أننا طبقنا صورتين من الاختبار أو أننا طبقنا اختباراً وقسمناه إلى قسمين ، وكان عدد المفحوصين ١٠ أفراد . وكانت درجاتهم فى مرتى القياس تحت العمود س والعمود ص . فيكون معامل الارتباط بين هذه الدرجات ، والذى رمزنا إليه بالرمز : $ر_{ق . ق}$

الأفراد	القياس الأول س	القياس الثانى ص	انحراف القياس الأول عن المتوسط ح س	$ح^٢$ س	انحراف القياس الثانى عن المتوسط ح ص	$ح^٢$ ص	ح س ح ص	الفرق بين الدرجات فى القياسين	مربع الفرق بين الدرجات فى القياسين
محمد	١٣	١١	+ ٥ر٥	٢٥ر٣٠	+ ٣	٩	+ ٥ر١٦	+ ٢	٤
على	١٢	١٤	+ ٥ر٤	٢٥ر٢٠	+ ٦	٣٦	+ ٢٧	− ٢	٤
حسن	١٠	١١	+ ٥ر٢	٢٥ر٦	+ ٣	٩	+ ٥ر٧	− ١	١
زكى	١٠	٧	+ ٥ر٢	٢٥ر٦	− ١	١	− ٥ر٢	+ ٣	٩
طه	٨	٩	+ ٥ر	٢٥ر	+ ١	١	+ ٥ر	صفر	صفر
فريد	٦	١١	− ٥ر١	٢٥ر٢	+ ٣	٩	− ٥ر٤	− ٥	٢٥
موسى	٦	٣	− ٥ر١	٢٥ر٢	− ٥	٢٥	+ ٥ر٧	+ ٣	٩
رضى	٥	٧	− ٥ر٢	٢٥ر٦	− ١	١	+ ٥ر٢	− ٢	٤
كرم	٣	٦	− ٥ر٤	٢٥ر٢٠	− ٢	٤	+ ٩	− ٣	٩
رفعت	٢	١	− ٥ر٥	٢٥ر٣٠	− ٧	٤٩	+ ٥ر٣٨	+ ١	١
المجموع	٧٥	٨٠	صفر	٥ر١٢٤	صفر	١٤٤	١٠٢		٦٦
المتوسط	٥ر٧	٨		مجـ $ح^٢$ س		مجـ $ح^٢$ ص	مجـ ح س ح ص		٦ر٦

الانحراف المعيارى لدرجات الأفراد فى القياس الأول :

$$ع_{س} = \sqrt{\frac{مج\, ح^{٢}_{س}}{ن}} = \sqrt{\frac{١٢٤{,}٥}{١٠}} = \sqrt{١٢{,}٤٥} = ٣{,}٥٣$$

الانحراف المعيارى لدرجات الأفراد فى القياس الثانى :

$$ع_{ص} = \sqrt{\frac{مج\, ح^{٢}_{ص}}{ن}} = \sqrt{\frac{١٤٤}{١٠}} = \sqrt{١٤{,}٤} = ٣{,}٧٩$$

معامل الثبات وهو معامل الارتباط بين درجات الأفراد فى القياسين :

$$ر_{س.ص} = \frac{مج\, ح_{س}.\, ح_{ص}}{ن\, ع_{س}\, ع_{ص}}$$ حيث ن هى عدد الأفراد (١٠)

$$= \frac{١٠٢}{١٠ \times ٣{,}٥٣ \times ٣{,}٧٩} = \frac{١٠٢}{١٣٣{,}٧٨٧} = +\ ٠{,}٧٦$$

ويمكن حساب معامل الارتباط هذا بدون حساب الانحرافات المعيارية وذلك باستخدام المعادلة التالية :

$$ر_{س.ص} = \frac{مج\, ح_{س}\, ح_{ص}}{\sqrt{(مج\, ح^{٢}_{س})\, (مج\, ح^{٢}_{ص})}} = \frac{١٠٢}{\sqrt{(١٢٤{,}٥)\, (١٤٤)}}$$

$$= \frac{١٠٢}{\sqrt{١٧٩٢٨}} = \frac{١٠٢}{١٣٣{,}٩} = +\ ٠{,}٧٦$$

بقى أن نحسب التباين الحقيقى لدرجات الأفراد فى القياسين من المعادلة التالية :

$$ر_{ق ق} = \frac{ع^2_{\infty}}{ع^2_{ق}}$$

حيث $ر_{ق ق}$ = ٠ر٧٦ وحيث $ع^2_{ق}$ = ٦ر٦

$$\therefore +٠ر٧٦ = \frac{ع^2_{\infty}}{٦ر٦}$$

$$\therefore ع^2_{\infty} = + ٠ر٧٦ \times ٦ر٦ = ٥ر٠٢$$

وهكذا اتضح أن معامل الثبات هو ٠ر٧٦ والتباين الحقيقى هو ٥ر٠٢ والتباين التجريبى هو ٦ر٦ وسنوضح فيما يلى مزايا وعيوب كل طريقة من طرق حساب معامل الثبات والمعادلات التى تعالج هذه العيوب .

طرق حساب معامل الثبات

قلنا إن للثبات نوعان ؛ ثبات مطلق وثبات نسبى . وعند حساب معامل الثبات نحصل على درجات الأفراد فى الاختبار بإجرائه أكثر من مرة ثم نحللها لنعرف مدى الاتساق بينها فى المرات المختلفة . ويحسن بالطبع أن تتكرر مرات الإجراء ، ولكننا نعلم أن سلوك الأفراد فى موقف الاختبار يتأثر بخبرتهم من هذا الموقف ، فهو يتأثر مثلا بالتدريب والخبرة والتذكر والنمو . ولهذا فبدلا من تكرار الإجراء نكتفى بالإجراء مرتين فقط على أن يكون عدد المفحوصين كبيراً وممثلا .

وقد عرفنا معامل الثبات بأنه معامل الارتباط بين مجموعتين من القياسات المتكافئة لخاصية معينة لدى مجموعة من الأفراد . وعلينا الآن أن نحدد هذه القياسات المتكافئة . وتنحصر الخطوات العملية فى تحديد القياسات المتكافئة ، وبالتالى فى تحديد معامل الثبات ، فى الطرق الآتية :

١ – إجراء صورتين متكافئتين من الاختبار وحساب معامل الارتباط بين درجات نفس الأفراد فى كل منهما Equivalent forms .

٢ – إجراء نفس الصورة من الاختبار على نفس الأفراد مرتين وحساب معامل الارتباط Test-Retest .

٣ – تقسيم الاختبار الواحد إلى مجموعتين متكافئتين من الفقرات ، وإعطاء درجة لكل فرد على كل من القسمين ، وحساب معامل الارتباط split-half .

٤ – تحليل التباين بين فقرات الاختبار ، وتحديد خطأ التباين Analysis of variance وهى أحياناً تسمى طريقة التكافؤ المنطقى -Rational Equivalence

٥ – معامل الثبات الحقيقى index of reliability

١ – معامل الثبات بطريقة الصور المتكافئة :

ما دمنا قد عرفنا معامل الثبات بأنه معامل الارتباط بين مجموعتين من القياسات المتكافئة ، فإن استخدام طريقة الصور المتكافئة يطابق هذا التعريف. ولكن علينا أن نحدد عمليات إعداد الصور المتكافئة حقاً من الاختبار . ويسمى جيليكسون هذه الطريقة بطريقة الصور المتوازية Parallel Forms ويسميها جيلفورد بطريقة الصور المتبادلة Alternate Forms ويسميها البعض الآخر بطريقة الصور المتكافئة Equinalent Forms .

ويفترض فى هذه الطريقة أنه يمكن إنشاء صورتين متكافئتين من الاختبار

الواحد ، متكافئتان من حيث تمثيل جانب السلوك المطلوب قياسه . أى أن التكافؤ يجب أن يشمل الجوانب الآتية :

(ا) عدد مكونات الوظيفة التى يقيسها الاختبار .

(ب) نسبة الفقرات التى تخص كلا منها .

(ج) مستوى صعوبة الفقرات .

(د) طريقة صياغة الفقرات .

(ه) طول الاختبار وطريقة إجرائه وتصحيحه وتوقيته :

(ز) تساوى متوسط وتباين ($ع^{2}$) درجات الأفراد على كل من الصور .

ويمكن الوصول إلى الصور المتكافئة من الاختبار الواحد بتقسيم عدد الفقرات ، التى اتضح أن مستوى صعوبتها يلائم الغرض ويناسب العينة ، بحسب تدرجها فى الصعوبة وبحسب تمثيلها للوظيفة المناسبة . وهكذا نصل إلى مجموعتين أو أكثر من الفقرات ذات الأوزان المتكافئة بحيث أن الفرد الواحد ، لو وحدت شروط الإجراء وكل العوامل المؤثرة على أدائه ، يحصل على نفس الدرجة تقريباً أيا كانت الصورة المطبقة عليه/ وقد يحتاج الباحث إلى عمل صورة مكافئة للاختبار الذى ثبتت صلاحيته . وهنا يقوم بصياغة فقرات كل فقرة منها تناظر فقرة من الاختبار الأصلى ، تناظرها من حيث مستوى الصعوبة ونوع الوظيفة المقاسة وطريقة الصياغة وطريقة الإجابة .

وبالرغم من أن فقرات الصور المتكافئة ليست متطابقة تماماً إلا أنه كلما اقتربت الصورة زادت فرصة انتقال أثر التدريب . فالمفحوص يمر بموقف الاختبار الأول على أنه موقف تعليمى وعلى هذا فهو يطبق ما اكتسبه من الإجراء الأول عندما يواجه موقفاً مشابهاً فى الإجراء الثانى فترتفع درجته

عند إعادة الإجراء بنفس الصورة أو بالصورة المكافئة . ولكن إذا أفاد كل المفحوصين من خبرتهم بالموقف الأول فى تنمية قدراتهم بنفس القدر فإنهم جميعاً سيحصلون على درجات تزيد بنفس المعدل عن درجاتهم الأولى ؛ فإذا تأكد الباحث من هذا كان معامل الثبات الذى يحسبه صحيحاً ، إذ أن المركز النسبي للفرد فى جماعته لن يتأثر ما دام الكل قد نمى بنفس القدر . وتتعرض هذه الطريقة لأخطاء تنشأ عن قصر أو طول المدة بين إجراء الصورتين المتكافئتين .

٢ – الثبات بطريقة إعادة الاختبار :

فى طريقة إعادة الاختبار Test—Retest Method يكرر تطبيق نفس الاختبار بنفس الصورة على مجموعة من المفحوصين . ومعامل الثبات هذا معرض لكثير من الأخطاء التى تقلل من أهميته . فإذا تأكدنا من أن إجراء الاختبار فى المرة الأولى لا يغير فى المفحوصين إطلاقاً ، بما يجعلهم يستجيبون فى الإجراء الثانى استجابات تمثل مستويات قدرتهم وحسب ، كنا محقين فى استخراج معامل الثبات بحساب معامل الارتباط بين درجات الأفراد فى المرتين . ولكن هذا لا يحدث ، إذ أن إجراء الاختبار على المفحوصين فى المرة الثانية يتأثر بالعوامل الآتية :

(ا) موقف الاختبار موقف تعيمى يفيد منه المفحوصون على الأقل فى الألفة بمواقف الاختبار التالية . فلا يكونون عند إعادة الإجراء متأثرين بغرابة الموقف والتوتر الانفعالى عند بدء أى امتحان بنفس الدرجة التى أثرت فى أدائهم لنفس الصورة من الاختبار فى المرة الأولى : وهذا النقص لا يؤثر وحسب على درجات الأفراد بل إنه يوجه إلى منهج البحث المتبع ، فلا نكون عند إعادة الإجراء قد وفرنا ووحدنا نفس ظروف الإجراء الأول ، وبالتالى فإن درجات الأفراد فى مرة تتأثر بعوامل معينة (كالاضطراب الانفعالى

وصدمة الجدّة وخطوات تكوين العلاقة الطيبة فى الإجراء الأول) ، بينما تمثل درجاتهم فى الإجراء الثانى محصلة عدة قوى منها الألفة بموقف الاختبار وتعلم الطريقة الاقتصادية فى معالجته .

(ب) وبالمثل ، ولنفس السبب ، لا تمثل درجات المفحوصين فى إعادة الإجراء مستويات قدرتهم وحسب بل وأيضاً تعد محصلة لعدة قوى منها انتقال أثر التدريب والمران Practice effect . كما أن هذه القوى تشمل أيضاً تعلم المفحوص طرقاً فنية فى الإجابة عن الفقرات ومعالجة مواد وأجهزة الاختبار .

(ج) هذا إلى أن الفترة بين الإجراء الأول والإجراء الثانى interval تؤثر هى الأخرى على أداء نفس الصورة من الاختبار عند إعادة الإجراء . ويمكن توضيح هذا التأثير فى النقاط التالية :

١ ـ إذا طالت هذه الفترة زاد احتمال تأثير النمو Development العقلى والجسمى والوجدانى والاجتماعى الحادث فيها لدى الأفراد . وتزيد خطورة تأثير النمو كلما صغر سن المفحوصين . فنحن نعلم أن معدل سرعة النمو فى الأعمار الصغيرة يكون كبيراً بما يجعل الاستجابة لنفس الصورة عند إعادة إجرائها بعد شهر من المرة الأولى انعكاساً لمقدار هذا النمو . ويقل معدل سرعة النمو كلما كبر سن المفحوصين ، وعلى هذا يزداد معامل الثبات بطريقة الإعادة فى حالة عينة من الكبار سناً عنه فى حالة الأصغر سناً .

٢ ـ ويزداد أثر عامل النمو كلما كان المفحوصون على مستوى عال من القدرة العقلية . فنحن نعلم أن معدل سرعة النمو يكون كبيراً عند الأذكياء عنه عند الأقل ذكاء . وعلى هذا يزداد معامل الثبات كلما انخفض ذكاء المفحوصين إذ أن النمو الحادث أقل مقدارا منه فى حالة الأذكى .

٣ – إذا قصرت الفترة بين الإجراء وإعادة الإجراء كانت استجابات المفحوصين فى الإعادة محصلة لعدة قوى ، فضلا عن تمثيلها لقدرتهم ، أهمها تذكر الأسئلة وحفظ الإجابات خاصة إذا كان الإختبار قصيرا أو كانت فقراته لا تتطلب إعمال وظائف معقدة .

وهكذا نجد أن طول الفترة بين الإجراء والإعادة يجب أن يتحدد بحسب طبيعة العينة من حيث الذكاء والسن ، وبحسب طبيعة الاختبار من حيث الطول وتعقد الوظائف المقاسة . وعموما نجد أنه ليست هناك قاعدة تحكم طول هذه الفترة ، وإن كانت غالبا لا تقل عن أسبوع ، فى الإختبارات الطويلة أو التى تقيس وظيفة معقدة ولا تزيد عن شهر فى الإختبارات القصيرة أو التى تقيس وظيفة بسيطة . ولكن هـذه المدة قد تكون أياما معدودة فى إختبارات الشخصية الطويلة مثل إختبار M.M.P.I.

وعلى هذا يجب أن يعمل حساب أثر هـذه المدة بين الإجراء والإعادة ، فإذا كان الأثر موحداً بين كل أفراد العينة ، كنا محقين فى استخدام هذه الطريقة إذ أن مركز الفرد بالنسبة لجماعته فى الوظيفة المقاسة لم يتغير كثيراً .

٣ – الثبات بطريقة نصفى الاختبار

تستدعى طريقة تجزئة الإختبار إلى نصفين Split – half method تقسيم الاختبار الواحد إلى قسمين يحصل كل فرد على درجة عن كل قسم منهما . وهكذا يصبح كل نصف وكأنه صورة مكافئة . إلا أننا فى الطريقة النصفية نطبق الاختبار كله (النصفين معا) ثم نصححه فيحصل الأفراد على مجموعتين من الدرجات إحداهما عن القسم الأول

والأخرى عن الثانى . بينما فى طريقة الصور المتكافئة تمر فترة بين إجراء الصورتين . فالفترة الزمنية بين إجراء القسمين ، فى الطريقة النصفية : تصبح صفرا . ولكى نضمن أن التقسيم لا يخضع للتبادل المنتظم Systematic alternation بين كل فقرة من القسم الأول مع نظيرتها من القسم الآخر ، نقسم الاختبار بحيث يصبح أحد القسمين حاويا للفقرات الفردية (الفقرة رقم ١ ؛ ٣ ، ٥ ، . . الخ) والقسم الآخر حاويا للفقرات الزوجية (٢ ، ٤ ؛ ٦ ، . . الخ) . وهكذا يوزع المفحوص جهده ووقته ويستغل وظائفه بنفس القدر فى الجزئين . بينما لا يتوفر هذا فى تقسيم الاختبار إلى جزئين الأول (الفقرات من ١ إلى ٥٠ مثلا) والأخير (الفقرات من ٥١ إلى ١٠٠ مثلا) إذ تؤثر عوامل تراكم التعب والملل فى الإجابة عن الجزء الأخير . فضلا عن أن القسم الأول عادة ما يكون أسهل من القسم الأخير .

والميزة المنهجية Methodological فى طريقة التجزئة النصفية إلى قسمين : فردى وزوجى ، هى إجراء النصفين فى تآنى Simultaneously فتوحد ظروف الإجراء توحيدا تاما . ولكن هذه الطريقة يعيبها أننا نحسب معامل الثبات من درجات الأفراد على كل نصف من الاختبار باعتباره اختبارا أو صورة كاملة مكافئة ، ولا نحسبه من درجات الأفراد عن الاختبار ككل . ونحن نعرف أن الطول الأصلى للاختبار يعنى درجة تمثيل عينة فقراته لجوانب الوظيفة المقاسة ، وعلى ذلك لا يشمل القسم كل منطقة السلوك المطلوب قياسه والاختبار ككل يمكن الاعتماد عليه ، لهذا السبب ، أكثر من نصفه . وهناك معادلات إحصائية تعوض هذا النقص .

(أ) معادلة سپيرمان براون :

هذه هى الصورة العامة لمعادلة سپيرمان براون Spearman-Brown Formula لإجراء تصحيح إحصائى لمعامل الثبات المحسوب بالطريقة النصفية :

$$r_{١١} = \frac{٢ r_{\frac{١}{٢} . \frac{١}{٢}}}{١ + r_{\frac{١}{٢} . \frac{١}{٢}}}$$

حيث $r_{١١}$ هى معامل ثبات الإختبار كله .

، $r_{\frac{١}{٢} . \frac{١}{٢}}$ هى معامل الارتباط بين درجات الأفراد على نصف الإختبار أى أنها معامل ثبات كل نصف من الإختبار . فلو كان معامل الإرتباط هذا هو ٦٨ر٠ يكون معامل الثبات للإختبار كله هو :

$$r_{١١} = \frac{٢ \times \text{٦٨ر٠}}{١ + \text{٦٨ر٠}} = \frac{\text{٣٦ر١}}{\text{٦٨ر١}} = \text{٨١ر٠}$$

ويعاب على هـــذه الطريقة أنها تتحدد بتساوى الانحراف المعيارى لكل من النصفين . فإذا لم يتساوى التباين بين الجزئين لجئنا إلى طريقة أخرى مثل .

(ب) معادلة جتمان :

تفيد هذه المعادلة فى حساب معامل الثبات للاختبار ككل ، من معامل ارتباط نصفية لو كانا غير متساويين فى تباين درجات الأفراد عليهما . والمعادلة هى :

$$ر_{١٠١} = ٢ \left(١ - \frac{ع_{ا}^{٢} + ع_{ب}^{٢}}{ع^{٢}} \right)$$

، $ر_{١٠١}$ هى معامل ثبات الاختبار ككل

، $ع_{ا}^{٢}$ هى مربع الانحراف المعيارى لدرجات الأفراد عن النصف ا أى تباينها .

، $ع_{ب}^{٢}$ هو تباين درجات النصف ب من نفس الاختبار .

فإذا كانت $ع_{ا}^{٢}$ (فى النصف ا) هى ٢ ، $ع_{ب}^{٢}$ (فى النصف) ب هى ٥ر١ ، $ع^{٢}$ للاختبار ككل هى ٦ كان معامل الثبات مساوياً :

$$ر_{١٠١} = ٢ \left(١ - \frac{٢ + ١٫٥}{٦} \right)$$

$$= ٢ \left(١ - \frac{٣٫٥}{٦} \right)$$

$$= ٢ \left(١ - ٠٫٥٨ \right)$$

$$= ٢ \times ٠٫٤٢$$

$$= ٠٫٨٤$$

(ج) معادلة رولون :

أما معادلة رولون Rulon Formula فهى تعتمد على ما افترضه صاحبها من أن تباين درجات الأفراد على الاختبار ككل يرجع ، كما سبق القول ،

إلى تباين حقيقى فى مستويات القدرة المقاسة كما يرجع إلى أخطاء تجريبية فى القياس . والمعادلة هى :

$$ر_{١٠١} = ١ - \frac{ع^٢_ف}{ع^٢}$$

حيث $ر_{١٠١}$ هى معامل ثبات الاختبار ككل ، $ع^٢_ف$ هى الخطأ التجريبى فى القياس وهو متوسط مربعات الفروق بين درجات الأفراد على نصفى الاختبار .

، $ع^٢$ تباين درجات المجموعة فى الاختبار ككل فيحصل كل فرد على درجتين إحداهما عن النصف الأول والأخرى عن النصف الثانى ، ونطرح الدرجتين من بعضهما بالنسبة لكل فرد ثم نربع الفروق ونجمعها ثم نقسمها على عدد الأفراد ثم نحسب متوسط درجات المجموعة على الاختبار ككل ثم نحسب إنحراف الأفراد عن هذا المتوسط ثم نربع هذه الانحرافات ونطبق المعادلة .

(د) معادلة هورست :

وضع هورست Horest معادلته للتغلب على صعوبات عملية تواجه الباحث عادة . وأهمها أنه قد يستحيل تقسيم الاختبار إلى نصفين متكافئين تماماً . فإذا فرضنا أن معامل الثبات للاختبار كله هو ($ر$) . وأن معامل الارتباط بين قسمى الاختبار هو ($ر_{س.ص}$) . وأن (س) هى نسبة عدد فقرات القسم الفردى إلى عدد فقرات الاختبار كله ، (ص) هى نسبة عدد فقرات القسم الزوجى إلى عدد فقرات الاختبار كله . يكون معامل الثبات هو :

$$\text{ر} = \text{ر}_{\text{س.ص}} \left[\sqrt{\text{ر}^{٢}_{\text{س.ص}} + ٤_{\text{س.ص}} \left(١ - \text{ر}^{٢}_{\text{س.ص}} \right) \text{ر}_{\text{س.ص}}} \right]_{\text{س.ص}} \Big/ ٢_{\text{س.ص}} \left(١ - \text{ر}^{٢}_{\text{س.ص}} \right)$$

٤ – معامل الثبات بطريقة تحليل التباين :

إن تقسيم الاختبار إلى نصفين إجراء مفتعل ، إذ أن الاختبار على قدر ما يحوى من الفقرات تكون هناك طرقاً لتقسيمه . كما أن حساب معامل الثبات من أحد قسمى الاختبار أو من القسمين معاً . فيه تجاهل لكثير من البيانات عن كل الفقرات التى يشملها الاختبار : فالثبات هنا هو اتساق الأداء من فقرة إلى أخرى فى الاختبار ، وهذا المفهوم قريب من مفهوم الاتساق الداخلى Internal Consistency . ولكن تحليل اتساق فقرات الاختبار يتأثر بالحدود الزمنية للإجراء . ولهذا فسنرى أن هناك معادلات لحساب ثبات اختبارات السرعة . وأهم معادلات حساب معامل الثبات بتحليل تباين (ع^{٢}) درجات الأفراد على فقرات الاختبار هى :

(ا) معادلة كودر – رتشاردسون :

تمتاز معادلة كودر – ريتشاردسون Kuder-Richardson Formula: بالبساطة العملية والسهولة والسرعة وصيغها هى :

$$\text{ر}_{١٠١} = \frac{\text{ن ع}^{٢} - \text{م} \left(\text{ن} - \text{م} \right)}{\left(\text{ن} - ١ \right) \text{ع}^{٢}}$$

حيث ر$_{١٠١}$ هى معامل ثبات الاختبار ككل ؛

وحيث ن هى عدد فقرات الاختبار : وليكن ٤٠ فقرة مثلا .

، ع٢ هى تباين درجات الأفراد على الاختبار . وليكن ٣٦ مثلا .

، م هى متوسط درجات الأفراد على الاختبار . وليكن ٢٥ر٧ مثلا .

وبالتعويض فى المعادلة نجد أن :

$$ر_{١١} = \frac{٤٠\times٣٦ - ٢٥ر٧ (٤٠ - ٢٥ر٧)}{٣٩\times٣٦}$$

$$= ٠ر٧٦$$

(ب) معادلة هويت :

تفترض طريقة هويت Hoyt's Analysis of Variance Procedure أن درجة الفرد على الاختبار قد تقسم إلى أربعة عناصر تكونها وهى :

١ – عوامل تتعلق بكل الفقرات عند كل الأفراد .

٢ – عوامل تتعلق بالفقرة .

٣ – عوامل تتعلق بالفرد .

٤ – عامل الخطأ .

وصورتها هى : معامل الثبات = ١ – $\frac{\text{خطأ التباين}}{\text{التباين بين الأفراد}}$

(ج) معادلة چلكسون :

تمتاز معادلة چلكسون Gulliksen عن معادلة سپيرمان براون بأنها تصلح لحساب معامل ثبات إختبارات السرعة ، التى يحدد فيها الزمن بحيث

لا يستطيع أى مفحوص أن يجيب عن كل الفقرات ، إذ أن عامل الزمن له أثره فى معامل الثبات كما سنرى . والمعادلة هى :

$$\bar{ر}_{١١} = ر_{١١} - \frac{م_{ن}}{ع^{٢}_{خ}}$$

حيث $\bar{ر}_{١١}$ هى معامل الثبات بعد تصحيح أثر السرعة .

، $ر_{١١}$ هى معامل الثبات قبل تصحيح أثر السرعة محسوب بطريقة سبيرمان براون وليكن ٩ر٠ ، $م_{ن}$ هى متوسط عدد الفقرات التى تركها الأفراد أى لم يسمح الوقت بمحاولة الإجابة عنها وليكن ١ فقرة .

، $ع^{٢}_{خ}$ هى تباين عدد الفقرات التى يجيب عنها الأفراد إجابة خاطئة وليكن ١٠ .

$$\therefore \bar{ر}_{١١} = ٩ر٠ - \frac{١}{١٠} = ٨ر٠$$

٥ ــ معامل الثبات الحقيقى :

معامل الثبات هو معامل الارتباط بين الاختبار ونفسه أو بين الدرجات الحقيقية على الاختبار ونفسه . وطريقة الثبات الحقيقى index of Reliability يعتمد على استخراج الجذر التربيعى لمعامل الثبات المستخرج بأى من الطرق الثلاث أى أن :

معامل الثبات الحقيقى = $\sqrt{\text{معامل الثبات المحسوب بأى معادلة}}$ فإذا كان معامل الثبات الذى حسبناه بأى معادلة هو ٦٤ر٠ كان معامل الثبات الحقيقى لهذا الاختبار هو $\sqrt{٦٤ر٠}$ = ٨ر٠ وهذا الرقم هو الحد الأعلى الذى لا تزيد عنه معاملات الارتباط بين الاختبار وأى محك آخر .

معادلة كرونباخ :

إشتق كرونباخ Cronbach صورة عامة لمعادلة معامل الثبات (على أساس معادلة كودر – ريتشردسن) وسماه المعامل الفا (α) وهى :

$$\text{معامل ألفا } (\alpha) = \frac{ن}{ن - ١} \left(١ - \frac{\text{مح ع}^{٢}\text{ك}}{\text{ع}^{٢}} \right)$$

حيث ن هى عدد أقسام الاختبار وهو غالباً ٢

، ك هو أحد أقسام الاحتبار

، مح $\text{ع}^{٢}$ك هو مجموع تباينات الأفراد فى هذا القسم

، $\text{ع}^{٢}$ هو تباين الاختبار الكلى .

تفسير معامل الثبات :

كل معامل ثبات حسبناه بإحدى الطرق أو المعادلات له معنى معين ولا يستخدم إلا فى حدود هذا المعنى . وكل معامل ثبات أوجدناه يقوم على شروط معينة يجب توفرها وإلا فلا يصح أن يستخدم . ومن هنا وجب عند ذكر معامل ثبات أى اختبار أن تذكر الطريقة والمعادلات المستخدمة والشروط التى توفرت قبل إيجاده . وبهذا فقط يمكننا أن نقارن بين معاملات الثبات بعـــد أن نعرف معنى كل معامل منها . ويقول كرونباخ Cronbach أن :

١ – معامل الارتباط بين مرتى الإجراء test-retest Method هو معامل استقرار Stability .

٢ – معامل الارتباط بين صورتى أو صور الاختبار المتكافئة equivalent

forms method هو معامل تكافؤ equivalence وهو يصبح معامل تكافؤ ومعامل استقرار معاً إذا مرت فترة زمنية مناسبة بين إجراء الصورتين .

٣ ــ أن معامل الارتباط بين درجات الأفراد على نصفى الاختبار split-half method هو معامل استقرار .

بمعنى أنه عند حساب معامل الثبات بطريقة الإعادة يهمنا أن نعرف مدى استقرار درجات الأفراد عن نفس الأسئلة بين مرات الإجراء . وعند حساب معامل الثبات بطريقة الصور المتكافئة يهمنا أن نعرف مدى ارتباط درجات الأفراد على صورتى أو صورة الاختبار أى مدى تكافؤ هذه الدرجات المعطاه عن صور متكافئة . بينما عند حساب معامل الثبات بالطريقة النصفية يهمنا أن نعرف التكافؤ ، بالمعنى [السابق ، بافتراض أن منطق الصور المتكافئة قائم بين الأنصاف المتكافئة .

وعلى هذا ، ما دمنا نجمع بين الاستقرار والتكافؤ فى طريقة تطبيق الصور المتكافئة بعد مرور فترة من الزمن بين الصورة والأخرى ، قد يتوقع البعض أن معامل الثبات بطريقة النصفين يكون أعلى منه بأى من الطريقتين الأخريتين . بينما نجد أن معامل الثبات النصفى دائماً أقل من معامل الاستقرار . وأقل من معامل التكافؤ أيضاً .

فمعامل التكافؤ يدل على مدى دقة قياس الاختبار لأداء الفرد فى لحظة معينة . فهو يقدر المدى الذى تتراوح فيه درجة الفرد إذا استخدمت عينات مختلفة من الأسئلة التى تقيس نفس القدرة . ومعامل التكافؤ يحسب عادة بالطريقة النصفية مصححة بمعادلة سبيرمان ــ براون .

وتعطى الطريقة النصفية نتائج مضللة إذا لم يكن النصفان متكافئين تماماً كالصور المتكافئة . فيتساوى المتوسط والانحراف المعيارى لكل من النصفين تساويهما بين كل من الصورتين . ويجب أن تتماثل الأنصاف

فى محتواها فلا يقسم اختبار الذكاء الذى يشمل فى نصفه أسئلة حساب وفى نصفه الآخر أسئلة أداء . وقد جرى تقليد أن يقسم الاختبار إلى نصفين أحدهما يشمل الفقرات الفردية والآخر يشمل الفقرات الزوجية . ولكن هذه الطريقة قد لا تؤدى إلى تكافؤ نصفى الاختبار . ويقترح كرونباخ أن تحلل أوراق الإجابة لتحديد مستوى صعوبة الفقرات عن طريق النسبة المئوية للناجحين فى كل فقرة ثم تقسم الفقرات إلى مجموعتين متكافئتين فى الصعوبة وفى المحتوى . ثم يحسب معامل الارتباط بين النصفين مهما كان ترتيب الأسئلة التى اعتبرت ضمن أحد النصفين .

وتقسيم الاختبار إلى نصفين متكافئين فى الصعوبة والمحتوى بالطريقة التى اقترحها كرونباخ يجعل معامل الثبات بالطريقة النصفية أكثر دقة . ولكن عدم الاهتمام بترتيب الفقرات التى يشملها كل نصف أمر يذهب بهذه الدقة . فقد يتضمن أحد النصفين ، بهذه الطريقة ، فقرات يجرى أغلبها فى النصف الأول من زمن الاختبار فى حين قد يتضمن النصف الآخر فقرات يجرى أغلبها فى النصف الثانى من زمن الاختبار . وهنا لا تكون كفاءة المفحوص متسقة إذ يتأثر فى النصف الآخر بتغير الظروف وبعوامل التعب والملل وانتقال أثر التدريب . ويمكن التغلب على هذا العيب بأن توزع الأسئلة ذات المستوى الموحد من الصعوبة والتى تتكافأ فى المحتوى توزيعاً متماثلا فى الاختبار بأن يجعل هذه الفقرات نصفين أحدهما فردى والآخر زوجى . ثم يحسب معامل الارتباط بين هذين النصفين فيكون معامل استقرار ومعامل تكافؤ معاً وفى نفس الوقت .

ومعامل التكافؤ يمكن أن يقاس أيضاً بطريقة الصور المتكافئة التى تدلنا مباشرة على ما إذا كانت العينات المختلفة من الفقرات تعطى نفس ا الصورة عن الفروق الفردية بين المفحوصين .

ومعامل الاستقرار يبين المدى الذى يتراوح فيه استقرار الدرجات ، عن فقرات اختبار معين ، فى مدى زمن معين هو الفترة بين الإجراء وإعادة الإجراء . فهو يدل على ما إذا كانت عينة السلوك المأخوذة فى وقت معين تطابق عينة السلوك المأخوذة فى أوقات أخرى : وهو يقدر بطريقة الإعادة فيأخذ نفس المفحوصين نفس الاختبار فى فرصتين مختلفتين ، بحيث لا تكون أمامهم فرصة للتدريب أو حفظ مواد الاختبار خلال الفترة الزمنية بين المرتين . ويعيب هذه الطريقة فضلا عن العيوب المذكورة سابقاً ، أنها لا تفيد فى اختبارات الاستدلال إذ قد يتذكر المفحوص حل المسألة فلا تصبح المسألة إذن مقياساً للاستدلال بل للتذكر .

وتقترب الطريقة النصفية كثيراً من معنى الاتساق فنحن نحسب اتساق أداء الفرد مع نفسه فى موقف الاختبار وينطبق هذا ، على وجه الأخص ، عند تعديل معامل الارتباط بين نصفى الاختبار باستخدام معادلة كودر — ريتشردسن التى تقيس الاتساق الداخلى بين الفقرات Inter-Item Consistency . وهكذا نقترب كثيراً من مفهوم التجانس Homogeneity . فيمكننا حساب معامل التجانس بطرح قيمة معامل كودر — ريتشردسن من معامل الثبات النصفى بطريقة سبيرمان — براون ؛ ذلك أن معامل الثبات بطريقة كودر — ريتشردسن يعتمد على تحليل التباين إذ يدخل فى اعتباره عدد الفقرات (ﻥ) والمتوسط الحسابى (م) والتباين (ع^{٢}) . وعموماً نجد أن معامل الثبات بطريقة كودر — ريتشردسون يعطى أصغر قيمة لقياس الثبات بينما تعطى معادلة سبيرمان — براون أعلى قيمة للثبات ويكون الفرق بينهما هو معامل التجانس .

وكما قلنا ليس هناك معامل ثبات كامل فالقليل من الاختبارات يقترب معامل ثباتها من واحد صحيح . وأحياناً تكون الاختبارات القصيرة ذات الثبات المنخفض ذات قيمة فى أغراض خاصة مثل إصدار الأحكام السريعة وعموماً ترتفع معاملات الثبات فى اختبارات القدرة والاستعداد عنها فى اختبارات الشخصية . وتختلف قيمة الثبات المطلوب باختلاف الغرض من استخدام الاختبار . ولكن أيا كان الغرض يحسن أن نصل إلى أعلى معامل ثبات ممكن . والجدول التالى (ص ٢٤٧) عن كرونباخ يقارن بين خصائص معاملات الثبات .

العوامل التى تؤثر على معامل الثبات .

رأينا أن العوامل التى تؤثر على ثبات الاختبار هى :

١ ــ طول الاختبار : فمن المعلوم أن الاختبار عينة للجانب من السلوك المطلوب قياسه . ولهذا فكلما كبر حجم هذه العينة ، أى كلما طال الاختبار زاد تمثيلها للمجتمع الأصلى أى لمنطقة السلوك المراد قياسها . وذلك لأن درجة الفرد فى الاختبار الطويل تمثل قدرته . وقدرته ثابتة ولهذا فدرجته هنا أكثر ثباتا . وعلى هذا يمكن رفع معامل الثبات بإطالة الاختبار Lengthening the test وذلك بالمعادلة الآتية :

$$\text{ر}'_{١٠١} = \frac{\text{ن}\ \text{ر}_{١٠١}}{١ + (\text{ن} - ١)\ \text{ر}_{١٠١}}$$

حيث $\text{ر}'_{١٠١}$ هو المعامل الذى نريد أن نصل إليه .

، $\text{ر}_{١٠١}$ هو المعامل الحالى .

، ن هى نسبة طول الاختبار الجديد إلى الاختبار الأصلى .

الاسم وطريقة الحساب	الفروض	الأخطاء المحتملة عند عدم توفر الفروض	تساؤلات يجيب عنها المعامل
معامل التكافؤ			
الطريقة النصفية	تكافؤ الأنصاف	ترتفع معاملات ثبات اختبارات السرعة عند استخدام هذه الطريقة بدرجة كاذبة . وينخفض المعامل بدرجة شـديدة إذا لم تتكافأ الأنصاف .	
كودر – ريتشردسون	يجب أن يقيس الاختبار عاملا واحدا	يرتفع معامل ثبات اختبارات السرعة بدرجة كاذبة عند استخدام هذه المعادلة بينما ينخفض هذا المعامل بدرجة كبيرة عندما تقيس الفقرات عوامل كثيرة .	ما مدى دقة الاختبار فى القياس ؟ ما مدى كفاءة تمثيل الاختبار لكل الفقرات التى كان يمكن أن يتضمنها ؟
الصور المتكافئة ، بحيث لا تمر فترة طويلة بين إجرائها .	يجب أن تتكافأ الصور	يبين المعامل درجة التكافؤ ، أكثر مما يبين مدى دقة قياس كل صورة للوظيفة المعينة .	
معامل الاستقرار إعادة الإجراء بعد مرور فترة	يجب أن لا تكون هناك فرصة لنمـو القدرة بفعل التدريب خلال الفترة بين الإجرائين .	تدريب الوظيفة المقاسة يخفض هذا المعامل .	ما مدى استقرار القياس بهذا الاختبار ؟
معامل الاستقرار والتكافؤ معاً إجراء الصور المكافئة بعد فترة من الزمن	يجب أن تتكافأ الصور ، وأن لا تكون هناك فرصة للتدريب	إذا لم يتوفر الفرضان انخفض تقدير هذا المعامل لدقة الاختبار فى القياس .	ما مدى تعلق عينة السلوك فى وقت بنتائج نفس العينة فى وقت آخر ؟

جدول (٣) يقارن بين خصائص وشروط ونقائص طرق حساب معاملات الثبات عن :

Cronbach, L. J., Eessentials of psycho'ocisl testing, N. Y. : Haper & Brothers Publishers, 1949.

فإذا أردنا رفع معامل ثبات الاختبار من ٥ر٠ إلى ٨ر٠

$$\therefore\ \text{٨ر٠} = \frac{\text{ن} \times \text{٥ر٠}}{\text{١} + (\text{ن} - \text{١})\ \text{٥ر٠}}$$

$$\text{٥ر٠ ن} = \text{٨ر٠} + \text{٤ر٠ن} - \text{٤ر٠}$$

$$\therefore\ \text{٥ر٠ ن} - \text{٤ر٠ ن} = \text{٤ر٠}$$

$$\therefore\ \text{١ر٠ ن} = \text{٤ر٠}$$

$$\therefore\ \text{ن} = \text{٤}$$

وهكذا يضاعف طول الاختبار إلى أربعة أمثال طوله الأصلى أى أن عدد الفقرات التى سنصممها أو نضيفها يساوى ٣ أمثال عدد فقرات الاختبار الأصلى : على أنه يجب ألا يطول الاختبار بحيث يصبح مملا :

(ب) زمن الاختبار :

معامل الثبات يظل يرتفع بزيادة الوقت الذى يستغرقة إجراء الاختبار ، إلى درجة معينة ، ثم ينخفض الثبات بطول المدة . وهذه الدرجة تختلف من اختبار إلى آخر .

(ح) تجانس المجموعة :

يقل ثبات الاختبار بزيادة تجانس homogeneity المجموعة المطبق عليها ، ويرتفع الثبات بزيادة تباين درجات أفرادها عليه . وعلى هذا يمكن التنبؤ بمدى تغير معامل الثبات بتغير الانحراف المعيارى لدرجات عينة جديدة من الأفراد باستخدام المعادلة التالية :

$$ر_{ص.ص} = ١ - \frac{ع^٢_{س}}{ع^٢_{ص}} (١ - ر_{س.س})$$

حيث $ر_{ص.ص}$ هى معامل الثبات عند إجراء الاختبار على المجموعة ص.

، $ع^٢_{س}$ هو تباين المجموعة س (وليكن ٦٢٥)

، $ع^٢_{ص}$ هو تباين المجموعة ص (وليكن ٤٠٠)

، $ر_{س.س}$ هو ثبات المجموعة س (وليكن ٩ر٠)

$$\therefore ر_{ص.ص} = ١ - \frac{٦٢٥}{٤٠٠} (١ - ٩ر٠) = ١ - \frac{٥ر٦٢}{٤٠٠}$$

$$= ١ - ١٦ر٠$$

$$= ٨٤ر$$

(د) الصدق :

يتوقف معامل الصدق على معامل الثبات فيزيد بزيادته وينخفض بانخفاضه . وكما رأينا فإن معامل الصدق لا يمكن أن يزيد عن الجذر التربيعى لمعامل الثبات . فإذا كان معامل ثبات أحد الاختبارات هو ٨١ر٠ ، فإنه لا يمكن إجراء أى تعديل يؤدى إلى أن يرتفع معامل الصدق ، فى حدود هذا الثبات ، عن $\sqrt{٨١ر٠} = ٩ر٠$. ولكننا رأينا أيضاً أن معامل ثبات الاختبار يتوقف على طوله ، فيزداد ثبات الاختبار بزيادة طوله إلى درجة معينة . فإن طال الاختبار أكثر مما يجب كان الجزء الأخير فيه متأثراً بالتعب والملل وتغير ظروف الإجراء . وما دام معامل الثبات يرتفع بإطالة الاختبار ، فإن معامل الصدق أيضاً يرتفع .

وعلى ذلك يمكننا إيجاد العلاقة بين معامل الصدق ومعامل الثبات عند إطالة الاختبار العدد (ن) من المرات (وليكن ٢) بفرض أن معامل ثبات الاختبار بطوله الأول ($ر_{\text{س . س}}$) هو (٠ر٧٠) وأن معامل صدق الاختبار قبل الإطالة ($ر_{\text{س . ص}}$) هو (٠ر٥٦) فيمكن حساب معامل ثبات الاختبار بعد مضاعفة طوله ، بتطبيق المعادلة :

$$ر_{\text{٢ س . ص}} = \frac{ر_{\text{س . ص}}}{\sqrt{\frac{\text{١} - ر_{\text{س . س}}}{\text{ن}} + ر_{\text{س . س}}}}$$

$$= \frac{\text{٠ر٥٦}}{\sqrt{\frac{\text{١} - \text{٠ر٧}}{\text{٢}} + \text{٠ر٧٠}}}$$

$$= \frac{\text{٠ر٥٦}}{\sqrt{\frac{\text{٠ر٣}}{\text{٢}} + \text{٠ر٧٠}}}$$

$$= \frac{\text{٠ر٥٦}}{\sqrt{\text{٠ر١٥} + \text{٠ر٧٠}}}$$

$$= \frac{\text{٠ر٥٦}}{\sqrt{\text{٠ر٨٥}}}$$

$$= \frac{\text{٠ر٥٦}}{\text{٠ر٩٢}}$$

$= \text{٠ر٦١}$ تقريباً .

(هـ) مستوى صعوبة الفقرات :

(و) خصائص العينة .

(ز) طريقة القياس

(ح) خصائص المقياس .

(ط) الخطأ التجريبي في القياس :

(ى) أخطاء الطرق الإحصائية :

(ك) التخمين والصدفة والغش والتزييف .

(ل) الانتباه والتذكر والحفظ والتعلم وأثر التدريب والنمو وأثر التعب والملل والألفة .

(م) المستوى الوظيفي للفرد وقت الإجراء أو إعادته ذلك من حيث الصحة والإنفعال ومستوى القدرة كما تتأثر بالشخصية :

(ن) طريقة حساب الثبات ومعادلاته .

بيانات الثبات :

يختلف معامل ثبات الاختبار من مجموعة إلى أخرى من الأفراد بل وبين المستويات المختلفة من العينة الواحدة . وهذه الفروق لا يمكن التنبؤ بها إحصائياً ولكن يمكن توقعها بالتجربة أي بتجريب الاختبار على عينات مختلفة لحساب ثباته . فالاختبار الواحد قد يقيس وظائف مختلفة إذا طبق على أفراد من مستويات مختلفة من حيث القدرة . كما أن السن والذكاء والتجانس كلها خواص في العينة تؤثر في الثبات وفي طريقة الإجابة عن الفقرات ومدى الاعتماد على التخمين وعامل الصدفة .

والخلاصة أنه عند ذكر معامل ثبات الاختبار يجب توضيح تفاصيل التجربة من حيث عدد وخصائص العينة والفترة بين القياسات المتتالية وأنواع معاملات الارتباط وطرق حساب الثبات . كما يجب أن يجرب الاختبار على عينات التقنين التي تنقسم بدورها إلى مجموعات فرعية كل

مجموعة منها متجانسة من حيث خصائص معينة . ويراعى عند استخدام الاختبار فيما بعد أن يطبق الاختبار على عينات مشابهة حتى يمكن وضع مستوى للثقة وتحديد نسبة متوقعة لثبات درجات الأفراد فيما بعد فإذا اختير الفرد بناء على درجته على الاختبار كان الاختيار سليما لأن درجته ثابتة . وإلا فإذا لم تكن الدرجة ثابتة بما يكفى ، أو كان خطأ التجريب فى القياس كبيراً ولم تكن الدرجة بالتالى ممثلة لمستوى القدرة الحقيقى لدى الفرد ، كان حكمنا مبنياً على قطاع صغير من السلوك ، قطاع لا يتميز بالثبات ولا بالشمول والجوهرية . ومن هنا تأتى أهمية حساب الثبات لا فى الاختبارات والمقاييس وحسب بل وأيضاً فى كل التقديرات والتقييمات التى نصدرها على الأفراد والجماعات .

الفصل الثالث

تحليـــل الفقرات

- مقدمة .
- اعتبارات في تحليل الفقرات .
- طرق تحليل الفقرات .
- الصدق المتبادل .
- ثبات الفقرات .
- وزن الإجابات عن الفقرات .
- استخدام صدق الفقرات .
- استخدام الاتساق الداخلى .

خاتمـــة .

الفصل الثامن

تحليل الفقرات

مقـــدمة :

درجة الفرد على المقياس هى مجموع أو محصلة درجاته على فقراته . فالفقرات فى الاختبار ، كالوحدات فى أى مقياس ، هى الجزء الذى تؤدى خصائصه إلى أن يتميز الاختبار بخصائص معينة . فمن ناحية الكم تكون الدرجة الكلية على الاختبار نتيجة للدرجات على فقراته وخاصة إذا كان من من نوع مقياس النقط Point Scale حيث تكون الدرجة هى مجموع عدد الفقرات التى أجاب عنها إجابة صحيحة . ومن ناحية الكيف تقوم دراسة نتائج تطبيق الاختبار ، على حالات فردية ، على تحليل طريقة الإجابة والأخطاء النوعية والأخطاء الثابتة فى الإجابة عن فقرات بذاتها .

ومن الناحية الإحصائية ، يعتمد حساب مختلف المعالم الإحصائية للاختبار على المعالم الإحصائية لفقراته ؛ فيتأثر المتوسط ، والانحراف المعيارى ، والتباين فى الاختبار بخصائص فقراته . ومن ناحية القياس ، يؤثر عدد ونوع ومستوى الفقرات على ثبات وصدق الاختبار .

وطبيعى أن الاختبارات المختلفة تحوى أنواعا مختلفة من الفقرات سواء من حيث الصياغة أو المواد أو المستوى . وقد عرضنا لهذا فى الفصل الخامس كما سنتناوله تطبيقيا فى قسم القياس التربوى . وعلى أى حال فإن تحليل الفقرات يمكن أن يعرف هكذا :

« تحليل الفقرات هو الدراسة التى تعتمد على التحليل المنطقى الإحصائى والتجريبى لوحدات الاختبار بغرض معرفة خصائصها وحذف أو تعديل

أو إبدال أو إضافة أو إعادة ترتيب هذه الفقرات حتى يتسنى الوصول إلى اختبار ثابت صادق مناسب من حيث الطول والصعوبة » .

اعتباران فى تحليل الفقرات :

بناء على ما تقدم نجد أننا نهتم فى تحليل الفقرات بناحيتين :

(ا) هى مستوى صعوبة Difficulty Level الفقرة بالنسبة لمن ستطبق عليهم وبالنسبة لموقعها فى الاختبار .

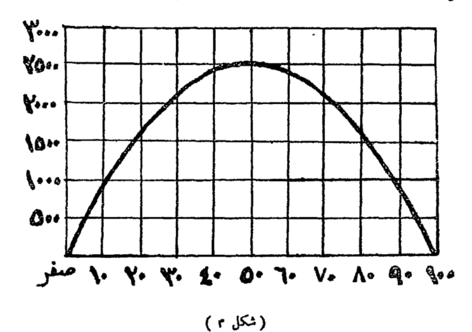
(ب) هى الدرجة التى تميز بها الفقرة بين مستويات قدرة المفحوصين . وفى الناحية الثانية ، أى قوة التميز Discriminative Power ، قد نقارن الأفراد على أساس أدائهم لمجموعة من الفقرات وهنا نهتم بالاتساق الداخلى internal consistancy بين الفقرات . وقد نقارن بين الأفراد على أساس محك خارجى هو أداء العمل وهنا نهتم بصدق كل فقرة Item Validity .

(ا) مستوى الصعوبة :

إذا كان للفقرة أن تميز بين الدرجة العالية وبين الدرجة المنخفضه من السمة المقصودة ، ما وجب أن تكون سهلة بدرجة أن ينجح فيها الجميع أو صعبة يفشل فيها الجميع . وعلى وجه العموم تحقق الفقرة الواحدة أقصى نجاح فى التميز بين الأفراد إذا كان مستوى صعوبتها يسمح بنجاح ٥٠٪ من أفراد المجموعة فى الإجابة عليها .

فإذا أخذنا فقرة واحدة ، أمكننا أن نتصور احتمالات قدرتها على التميز فيما يلى . فإذا أعطيت الفقرة لمفحوصين عددهم ١٠٠ ونجح فيها فرد واحد كانت تميزه عن الباقين وعددهم ٩٩ . ويكون حاصل الضرب ١ × ٩٩ = ٩٩ . أما إذا نجح فيها ١٠ أفراد وفشل ٩٠ كانت الفقرة تميز بين كل فرد من الناجحين العشرة وكل فرد آخر من جماعة الفاشلين التسعين . فإذا طبقت الفقرة على الأفراد فردين فردين كان عدد الحالات التى تنجح

فيها الفقرة فى تمييز فردى الثنائيات هو ١٠ × ٩٠ = ٩٠٠ حالة ؛ أما إذا نجح فيها ٥٠ فرداً وفشل ٥٠ فرداً فإنها تميز بين كل عضو فى مجموعة الخمسين الناجحين وكل عضو آخر فى مجموعة الخمسين الفاشلين ؛ فإذا أعطيت الفقرة لكل أفراد المجموعة اثنين فاثنين مع التبديل نجحت فى تمييز فردى الثنائى فى ٥٠ × ٥٠ = ٢٥٠٠ حالة . ويبين الشكل التالى (شكل ٣) العلاقة بين مستوى الصعوبة ، معبرا عنه بالنسبة المئوية الناجحة فى الإجابة على الفقرة وبين عدد المرات التى تنجح فيها الفقرة فى التمييز بين الثنائيات .



(شكل ٣)

واضح إذن أن الفقرة التى ينجح فيها ٥٠٪ من الأفراد تميز ذا القدرة العالية عن ذى القدرة المنخفضة فى ثنائيات الأفراد فى عدد من هذه الثنائيات يفوق ما تستطيعه الفقرة الصعبة جدا أو السهلة جداً .

وتشير أغلب الدراسات النظرية والعملية إلى أنه يحسن الاعتماد على الاختبار الذى يميز بين أعضاء الجماعة المختبرة عندما يسمح متوسط مستوى صعوبة فقراته بنجاح ٥٠٪ تقريبا من الجماعة فى كل فقرة : والتوزيع الصحيح لصعوبة الفقرات يضع فى الاعتبار أيضاً العلاقة بينها : ففى حالة

ارتباط كل فقرة بكل فقرة أخرى ارتباطا عالياً يتطلب الإختبار تشتتا أكثر فى مدى صعوبة فقراته حول مستوى الصعوبة المتوسط ٥٠ ٪. أما إذا انخفض معامل الإرتباط بين كل فقرة والأخرى فيحسن أن يشمل توزيع صعوبة الفقرات مدى أصغر ، ولكننا لا نستطيع وضع توزيع أمثل لكل إختبار بحسب معاملات الإرتباط الداخلية بين فقراته. إلا أن هناك اتفاقا عاما على أن الفقرة الصالحة متوسط مستوى صعوبتها يجب أن يسمح بنجاح ٥٠ ٪ من الأفراد عليها كما يجب أن توضح درجة من تشتت استجاباتهم عنها .

(ب) القدرة على التمييز :

نعنى بالقدرة على التمييز قدرة الفقرة على أن تميز بين الأفراد الحاصلين على درجات مرتفعة وبين من يحصلون على درجات منخفضة فى السمة التى تقيسها الفقرات كلها (أى الإختبار) . وهذا التمييز قد يكون على أساس الدرجة الكلية على الإختبار ، أوعلى أساس درجة على محك خارجى كالنجاح فى الدراسة أو العمل . وهذا الأساس يختلف من حالة إلى أخرى تبعا للغرض من إجراء الإختبار . وأحيانا نجمع الأساسين معاً . ولكى نوضح المواقف التى تدعو إلى تحليل العلاقة بين الفقرة وبين الدرجة على الإختبار ككل (وتسمى هذه العملية بتحليل الإتساق الداخلى internal consistency analysis) والمواقف التى تدعو إلى تحليل العلاقة بين الفقرة وبين المحك الخارجى external Criterion (وتسمى هذه العملية بإيجاد صدق الفقرة item Validation) فإننا نعرض الآن لنوعين من أدوات القياس .

والنوع الأول من أدوات القياس نعطى فيه درجة واحدة كلية على الإختبار . والدرجة الكلية على الإختبار لا تعنى أن هناك وحدة أساسية فى الفقرات نفسها ، ولكنها تعنى الإعتبارات العملية فى استخدام الإختبار .

وأوضح مثال على هذا النوع من أدوات القياس نجده فى استمارة البيانات عن التاريخ الشخصى Bibliographical Data Blank التى تشمل أسئلة جد مختلفة عن التاريخ الماضى للفرد . ويشمل هذا النوع أيضاً استبيانات الشخصية Personality Questionnaires وقوائم الميول Interest Inventories . وعند إعداد هذه الأدوات يجعلها الباحث تشمل مدى واسعاً من الفقرات تكون العلاقات بينها سطحية . فهى ليست علاقة وظيفية أساسية ، ولكنها علاقة سطحية أتت من حقيقة أن كل الفقرات تشير إلى حقائق عن التاريخ الشخصى ، أو عبارات تتعلق بما يحبه الفرد وما لا يحبه ، أو عبارات عن المشاعر الفردية أو سلوك الفرد . فالسهولة العملية والتشابه السطحى فى صياغة الفقرات يدفعان إلى جمع هذا العديد المتنوع من الفقرات فى استمارة اختبار واحدة . وعلى هذا لا تكون هناك أسباب تجعلنا نعتقد أن فقرات هذا النوع تمثل جانباً موحداً مفرداً من النمو الفردى .

فى مثل هذا الموقف يهمنا أن نتساءل عن صدق كل فقرة مفردة . فكل فقرة فى ذاتها اختبار صغير مفرد . وفى الحكم على صلاحيتها يجب أن نهتم بصدقها . وليس هناك أساس مقدماً يحددنا فى تصحيح الفقرات وجمعها فى درجة واحدة . والأساس الوحيد فى انتقاء الفقرات أو إعطاء درجة عنها هو الصدق العملى Emperical Validiiy لكل فقرة . ولهذا السبب يكون من المناسب ، فى هذا النوع من الإختبارات ، أن نحدد قدرة كل فقرة (وقدرة كل استجابة محتملة ، بالنسبة لكل فقرة) على التمييز بين الأفراد على أساس محك خارجى كالنجاح فى الدراسة أو العمل . وتكون الإستجابة على الفقرة صادقة إذا كان الأفراد الذين استجابوا بها أكثر نجاحاً فيما يتعلق بالمحك الخارجى ، أو كان أغلبهم أكثر نجاحاً ،

من الأفراد الذين لم يختاروا هـــذه الإستجابة لتلك الفقرة .

وعندما تمثل فقرات الاختبار المنفصلة مجالا غير متجانس ، من السلوك ، لا يكون هناك سبب لتوقع الإرتباط بينها مقدماً . ولا يكون هناك ، بالتالى ، مبرر لتحليل علاقة الفقرة المفردة بالدرجة المعطاة عن مجموعة من الفقرات . وهنا لا نجد أساساً يمكننا من إعطاء درجة واحدة عن هذه المجموعة من الفقرات ، وبالتالى ، لا تكون للدرجة دلالة تدفعنا إلى بحثها بدراسة الإتساق الداخلى بغرض تحليل الفقرات .

أما النوع الثانى من أدوات القياس فيقوم على أساس قياس قطاع محدد معين من القدرة . هذا القطاع ، مثلا ، قد يكون مدى المفردات . وفى مثل هذا الاختبار ، تهدف كل فقرة إلى الإسهام فى تقدير جانب موحد مفرد من القدرة . وفى هذا النوع من الاختبارات أيضاً يسهل التعرف مقدماً على الإجابة الصحيحة عن كل فقرة ، بل ويسهل أيضاً تصحيح الفقرات دون الرجوع إلى محك خارجى فإذا أعد الباحث ١٠٠ فقرة من هذا النوع ، أمكنه أن يعد مقدماً مفتاح التصحيح للاختبار دون أن ينتظر ما تأتى به نتائج التطبيق . وفى هذه الحالة علينا أن نتساءل عن درجة قياس الفقرة للوظيفة التى يقيسها الإختبار ككل . وقد يكون هذا ممكناً لأننا (١) نحصل على درجة عن الاختبار ككل (٢) ولأن هذه الدرجة لها دلالة لأن واضع الإختبار كان يهدف إلى أن تسهم كل فقرة فى قياس وظيفة واحدة متجانسة .

وعندما تكون فقرات الإختبار متجانسة بمعنى أن كل فقرة تقيس نفس العوامل أو الجوانب من قدرة الفرد بنفس تنظيمها يفقد إرتباط الفقرات المفردة بالمحك الخارجى دلالته . فما دامت الفقرة المفردة تقيس نفس الجزء من القدرة تماماً ، فإن صدقها يختلف وحسب باختلاف ثباتها : بمعنى أن

معامل ارتباطها بالمحك الخارجي يعتمد فقط على معامل ثباتها أى ارتباطها بنفسها . وعندما تكون الفقرات متجانسة homogeneous حقاً ، ينعكس الفرق بين ثبات الفقرات على الفرق بين ارتباط هذه الفقرات بالدرجة الكلية على الإختبار . وفى هذه الحالة ، تجمع كل البيانات المتعلقة بالفقرات بواسطة القيام بتحليل الفقرات بالإتساق الداخلى ، ويكون تحليل الفقرات المفردة بربطها بمحك خارجي أمراً مضيعاً .

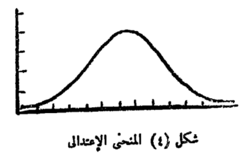
هذان النوعان من أدوات القياس طرفان لامتداد تقع على نقاطه المتفرقة أنواع أخرى متعددة . وأول الطرفين نجد الفقرة فيه يمكن أن تعد فى ذاتها إختباراً كاملا صغيراً ، وليست لها علاقة جوهرية بالفقرات الأخرى ويحسب صدقها بتقييم مزاياها . وعلى الطرف الآخر نرى إختباراً يتكون من فقرات كلها مقاييس متجانسة تماماً لنفس الوظائف فى الفرد . وفى الإختبارات من النوع الأول تتم عملية تحليل الفقرات المفردة وحسب بإيجاد صدق كل فقرة . وفى الإختبارات من النوع الثانى تتم عملية تقدير كل فقرة عند القيام بعملية تحليل الإتساق الداخلى . وكما قلنا ، لا نجد فى القياس حالتين متطرفتين تنطبق كل منهما على أحد النوعين ، فلكل إختبار درجة من التجانس ودرجة من التباين . وقد يسهل تصحيح الفقرة مقدماً كما قد يسهل الوصول إلى درجة كلية ، إلا أن الوظائف التى تقيسها الفقرات قد تكون متنوعة بدرجة تلزمنا بإيجاد صدق كل فقرة منها . ومن الجدير بالذكر أن نلاحظ أنه كلما زادت أهمية إيجاد صدق الفقرة المفردة قلت أهمية تحليل الإتساق الداخلى وكلما انخفضت دلالة صدق الفقرة المفردة زادت دلالة الإتساق الداخلى .

القياس الموضوعي للصعوبة :

قلنا إن الإختبار مقياس موضوعي . ولكن كيف يكون ذلك ؟ في عرضنا السابق ، اتضح ، أن التطبيق ، والتصحيح ، وتفسير الدرجات كل هذا موضوعي بمعنى أنها مستقلة عن الحكم الذاتي للفاحص نفسه . فدرجة المفحوص لا تتأثر بشخص الفاحص . هذا يحدث في حدود معينة ، ما دمنا لم نتوصل إلى تقنين كامل ولا إلى موضوعية تامة : ولكن على الأقل ، هذه الموضوعية هي هدف تكوين الاختبار ، وقد حصلتها أغلب الإختبارات ، بدرجة كبيرة ، غالبا .

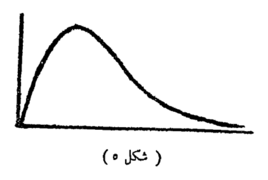
وهناك طرق هامة أخرى يمكن أن نحكم بواسطتها على موضوعية الإختبار بدقة فتحديد مستوى صعوبة الفقرة ، أو الإختبار كله ، وقياس ثبات الإختبار وصدقه ، مبنية على الموضوعية ، وهي خطوات عملية .

وليس ترتيب الفقرات وحسب ، بل وأيضاً اختيارها ، يمكن أن يعتمد على نسب من أجابوا عنها من العينات التجريبية . فإذا كانت عينة التقنين قطاعاً مستعرضاً ممثلا من المجموع السكاني المراد قياسه ، كان من المتوقع أن ترسم الدرجات في توزيعها منحنياً اعتدالياً . أي أن الأغلبية في الوسط . وعلى الطرفين أقليتين . والانحدار من القمة إلى الذيل يميناً ويساراً تدريجي انظر الشكل (٤) ولا يمكن الحصول على المنحنى الاعتدالي إلا من عينات ضخمة ممثلة .



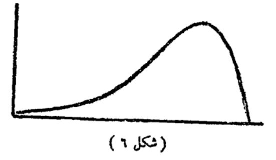
شكل (٤) المنحنى الإعتدالي

ولكن التواء المنحنى كما فى شكل (٥) ، وهو منحنى موجب الإلتواء Positively skewed فيعنى أن « أرضية » Floor الإختبار مرتفعة جداً بالنسبة لمن يراد تطبيقه عليهم ، وينقصه العدد الكافى من الفقرات السهلة التى تميز بدقة النهاية المنخفضة من المدى . وعلى ذلك فالأفراد المتوسطون الموزعون حول المتوسط سيحصلون على صفر أو درجات قريبة من الصفر :



(شكل ٥)

أما شكل (٦) وهو منحنى سالب الإلتواء Negatively skewed, فيعنى أن « سقف » Ceiling الاختبار منخفض فقد يحصل على الدرجات النهائية فيه بعض المفحوصين ولا يمكننا أن نعرف به أيهم أفضل أو أيهم يكون أداؤه أفضل لو كانت الفقرات أصعب :

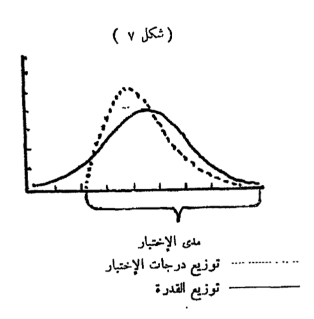


(شكل ٦)

وبعد رسم المنحنى يمكن أن نعدله بحيث يقرب من الإعتدال ذلك أنه لو كان ملتوياً نحو السهولة أمكننا إضافة بعض الفقرات الصعبة وحذفنا بعض السهلة وعدلنا البعض الآخر . ونفس الشيء ، بالعكس ، لو كان صعباً . على أنه أيضاً يمكن علاج الأمر بإعطاء وزن أكبر للفقرات الأصعب وأصغر للأسهل . وهكذا يمكن أن نجد نسبة كبرى من المفحوصين تستطيع الحصول على درجات عن ٥٠ ٪ من الفقرات . وهكذا يمكن الوصول إلى أقصى حد للتفريق بين الأفراد فى كل مستويات القدرة

وشكل (٧) يدل على أن الفقرات لاتقيس كل مدى القدرة ولا تصلح لتقييم المتطرفين من ضعاف القدرة أو المتفوقين فيها . كما أن قمة الإختبار تعلو قمة القدرة أى أن الإختبار يقيس مستوى بسيطاً من القدرة

(شكل ٧)



طرق تحليل الفقرات

(أ) دلائل صعوبة الفقرة :

نتعرف على مستوى صعوبة الفقرة من النسبة المئوية من الأفراد ، فى مجموعة معينة ، التى تستطيع أن تجيب عن السؤال أو أن تحل المشكلة :

وكلما صغرت النسبة المئوية الناجحة فى الفقرة ، زادت صعوبة الفقرة والعكس صحيح . وقد انتقد هذا الدليل على أساس أن طبيعة البيانات تتحدد بما تطلبه الإجابة عن الفقرة وليس تبعاً لتعقد الوظيفة العقلية التى تعمل فى الإجابة عن هذه الفقرة . فليست هناك صعوبة فى الوظيفة العقلية المستخدمة فى معرفة أن ٩ × ٨ = ٧٢ أكثر مما هو الحال فى معرفة أن ٣ × ٤ = ١٢ ؛ كما أن الوظيفة العقلية الأساسية المستخدمة فى الإجابة عن السؤال « ما هو الكنغر ؟ » بأنه حيوان ليست أصعب من الوظيفة المستخدمة فى الإجابة عن السؤال « ما هو الجمل ؟ » بأنه حيوان . إذ أن الصعوبة فى السؤال الأول متوقفة على ندرة وجود هذا الحيوان فى بيئتنا لا على تعقد المفهوم الذى نقيسه فى الحالتين . وهكذا نجد أن من يجيب على أحد السؤالين لا يقل فى الوظيفة المقاسة عمن يجيب عن السؤالين إذ أن هذا يرجع إلى صغر محصول خبرة الفرد وهو أمر لا علاقة له بالذكاء مثلا . وعلى هذا فإن نسبة النجاح فى الفقرة لا تعبر تماماً عن مستوى صعوبتها ، إلا أننا ما زلنا نقبل هذا المبدأ حتى نجد دليلا أفضل .

فالإعتماد على النسبة المئوية للناجحين فى الفقرة ، من بين المجموع السكانى الذى طبقت عليه ، كدليل على مستوى صوبتها أمر تحوطه عقبتان ، فربما كان بعض الناجحين فى إجابة الفقرة قد أجاب بالصـــدفة ، بالتخمين ، أو بطريقة لا تدل على زيادة معلومات الفرد أو على تفكيره فى المشكلة المعطاة له . والعيب الثانى هو أننا لو تصورنا أن القدرة التى يقيسها الإختبار موزعة اعتدالياً بين الأفراد ، كانت الفروق المتساوية بين النسب ، المئوية للنجاح غير متعلقة بدرجات مكافئة من الصعوبة أى أنه لا يمكن تمثيل العلاقة بين نسب الناجحين ودرجات الصعوبة بخط مستقيم . وسنعالج التعديلات التى يمكن إدخالها على هذه الطريقة للتغلب على هذين النقصين .

وأول نقص سنعالجه هو تدخل عوامل الصدفة محاولين تقدير النسب المئوية لأفراد الجماعة الذين توصلوا إلى الإجابة الصحيحة خلال معرفة صحيحة أو استدلال صحيح وأن نستبعد الذين أجابوا بالتخمين **guessing** . وأى تصحيح للتخمين يتضمن شيئاً من التقريب . وهذا التقريب يعتمد على افتراضين . والإفتراض الأول هو أن الفرد الذى يختار إجابة غير صحيحة عن الفقرة تنقصه المعلومات الصحيحة أو الفهم السليم وليس السبب هو سوء الفهم أو خلط المعلومات . والإفتراض الثانى هو أن الشخص الذى لا يعرف الإجابة الصحيحة من بين عدة إجابات مرفقة بالفقرة تجذبه كل هذه الإجابات ليختار واحدة منها بنفس الدرجة من الجذب فهو لا يعرف أيها الصحيح . إذا قبلنا هذين الفرضين ، توقعنا وجود نسبة معينة من الأفرد الذين لا يعرفون الإجابة الصحيحة فيختارون بمحض الصدفة إجابة يحدث أن تكون صحيحة . هذه النسبة المئوية من الأفراد هى متوسط النسب المئوية لمجموع الأفراد الذين يختارون كل إجابة من الإجابات الخاطئة . ونحن نستبعد هذه النسبة المئوية من النسبة المئوية للناجحين فى الفقرة لنصل إلى تقديرنا للنسبة المئوية التى عرفت حقاً الإجابة الصحيحة عن الفقرة . فتكون المعادلة هى :

$$\text{النسبة المئوية} = \frac{\text{ح} - \frac{\text{خ}}{\text{هـ} - ١}}{\text{ح} + \text{خ} + \text{ل}}$$

حيث النسبة المئوية هى النسبة المئوية للأفراد الذين يعرفون الإجابة الصحيحة عن الفقرة فعلا ،

، ح هى عدد من أجابوا إجابة صحيحة عن الفقرة .

، خ عدد من أجابوا إجابة خاطئة عن الفقرة نفسها .

، ك هو عدد من تركوا هذه الفقرة دون أى إجابة برغم أن الوقت سمح لهم بقراءتها .

، ن هو عدد متغيرات الاستجابة عن نفس الفقرة .

مثال :

أعطيت إحـــدى الفقرات لمفحوصين عددهم ٢٠ فرداً ، فأجاب ١١ فرداً منهم إجابة صحيحة وأجاب ٥ أفراد إجابة خاطئة ، وكان الوقت كافياً ولكن بعض المفحوصين قرأوا الفقرة ولم يعرفوا الإجابة إذ ترك ٤ أفراد الفقرة دون إجابة علماً بأن الفقرة متبوعة بعدد ٣ متغيرات يختار من بينها الفرد ، ما هى النسبة المئوية لمن يعرفون الإجابة الصحيحة فعلا عن هـــذه الفقرة بإعتبار النسبة دليل على مستوى سهولة الفقرة .

الإجابة :

النسبة المئوية لمن أجابوا بناء على معرفة صحيحة

$$= \frac{١١ - \frac{٥}{٣ - ١}}{١١ + ٥ + ٤} = \frac{١١ - \frac{٥}{٢}}{٢٠} = \frac{٨ر٥}{٢٠}$$

$$= ٠ر٤٢٥ = ٤٣٪ \text{ تقريباً .}$$

وبناء على هذا المفهوم لايكون هناك غموض فى المعادلة إلا الرمز ك الدال على عدد من اختاروا أن يتركوا الفقرة ؛ أى أن المفحوص هنا قرأ

الفقرة ثم تركها مختاراً . وفى إختبار السرعة أو الإختبار الموقوت Timed or speeded قد لا يفرق البعض بين الفقرات التى تركها المفحوصون مختارين لأنهم لم يعرفوا الإجابة عنها وبين الفقرات التى لم يصلوا إليها لضيق الوقت .

وهناك أساس عملى للتمييز بين الحالتين ، هو أن نفترض أن كل فرد وصل إلى فقرة معينة فى إجابته عن الإختبار قد حاول كل الفقرات التى قبلها والتى تليها مباشرة ، أما الفقرات ابتداء من التى تلى آخر فقرة أجاب عنها فنفترض أن الوقت لم يسمح بالإجابة عنها . وعلى هذا تكون الفقرات المتروكة قبل آخر فقرة أجاب عنها المفحوص هى فقرات لم يستطع إجابتها أو لم يعرف إجابتها . ونحن نرمز للفقرة الواحدة منها بالرمز ك .

فإذا كان الإختبار يتكون من ١٠٠ فقرة وأجاب أحد الأفراد عنه حتى وصل إلى الفقرة رقم ٧٠ فإننا نفترض أن الفقرات من ٧٢ حتى ١٠٠ ، وهى ٢٩ فقرة ، لم يتسع الوقت لحلها وتكون الفقرات من ٧١ حتى ١ كلها فقرات حاول إجابتها فإذا وجدنا بينها ٣ فقرات متروكة إفترضنا أنها أصعب من مستواه فى الوظيفة المقاسة فإذا تبين لنا أن ٢٠ فرداً من إحدى المجموعات تركوا إحدى الفقرات بنفس الطريقة كانت قيمة ك فى المعادلة هى ٢٠ .

يجب أن نلاحظ أن المعادلة السابقة تعتمد على الافتراضين معاً . فإذا كانت إحدى متغيرات الإجابة التى تتبع الفقرة المعينة متميزة عن بقية المتغيرات اختلف مقدار جذبها ، للأفراد الذين لا يعرفون الإجابة الصحيحة ، عن مقدار جذب غيرها . يحدث هذا سواء كانت واضحة الصواب إذ أن احتمال أن يختارها الفرد ، بطريق الصدفة أو التخمين ، الذى لا يعرف الإجابة الصحيحة يزيد عن احتمال أن يختار غيرها . ويحدث نفس الشيء لو كانت المتغيرة واضحة الخطأ إذ يقل احتمال أن يختارها نفس الفرد برغم أنه لا يعرف الإجابة الصحيحة . كما أن مقدار جذب كل متغيرة يتوقف على

عدد متغيرات الإجابة التى يجب على المفحوص أن يختار بينها : فإحتمال اختيار إحدى ٣ متغيرات يزيد عن احتمال إختيار إحدى ٥ متغيرات : وعلى هذا يجب أن نوحد عدد الإختيارات أمام الفرد فى كل الفقرات بقدر الإمكان كما يجب أن يتساوى جذب كل متغيرة منها :

ومن ناحية أخرى ، نجد أن الأفراد الذين يجيبون عن الفقرة إجابات خاطئة لا يفعلون هذا نتيجة لجهلهم وخطئهم فى التخمين بل نتيجة لعدم صحة معلوماتهم أو لخطأ العملية العقلية المستخدمة . فإذا كانت أسباب خطأ الأفراد فى الإجابة كلها أسباب أصيلة لما تدخلت الصدفة أو التخمين . وفى هذه الحالة أيضاً يكون كل الأفراد الذين أجابوا إجابة صحيحة قد فعلوا هذا لأن معلوماتهم صحيحة ولأنهم استخدموا عمليات عقلية سليمة . هنا لا يكون محل للصدفة ، ولا يكون هناك مجال لهذه المعادلة . وفى الواقع نحن لا نعلم إلى أى مدى يتدخل هذان العاملان (العامل الأول هو تدخل عامل التخمين فى كل من الإجابات الصحيحة وغير الصحيحة ، والعامل الثانى هو الجهل أو المعرفة وسوء استخدام العمليات العقلية أو سلامة التفكير كأسباب أصيلة تؤدى إلى الإجابات الخاطئة أو الصحيحة) . ولهذا فنحن نفترض أن العاملين يعوضان أثر بعضهما فتكون النتيجة هى أن النسبة المئوية الناجحة للفقرة هى النسبة المئوية من أفراد المجموعة التى عرفت الإجابة الحقة أو التى حلت المشكلة .

وبالرغم من أن القيمة المصححة بالنسبة المئوية التى عرفت الفقرة هى قيمة تقريبية وهى تتحدد بالافتراضات التى ناقشناها ، إلا أنها تعطى صورة أصدق من القيمة الخام . وهى طريقة تفيد فى مقارنة مستوى صعوبة الفقرات التى (١) تختلف اختلافاً واضحاً فى تكرار عدد المفحوصين الذين تركوها لصعوبتها . (٢) وبالنسبة للفقرات التى تختلف فى عدد المتغيرات وفى مقدار جذب كل منها : فالنسبة المئوية الخام لمن أجابوا إجابة صحيحة لا يمكن مقارنتها فى هذه الحالات :

ويمكن التغلب على العيب الثانى من عيوب النسبة المئوية ، وهو عدم تساوى الوحدات ، (انظر الفصل الثانى والتاسع) للناجحين فى الفقرة بإعتبارها دليلا على مستوى الصعوبة إذا افترضنا بعض الفروض المتعلقة بشكل توزيع السمة المقاسة فى المجموعة المختبرة . فالافتراض العادى أن السمة موزعة اعتداليا . ومن هنا تكون النسب المئوية ممثلة فى قيم على المقياس على المحور الأفقى الذى رسمنا عليه المنحنى الاعتدالى . وعلى هذا ، تكون نسبة النجاح ٥٠٪ مقابلة للقيمة صفر على المقياس ، ونسبة النجاح ٧٥٪ مقابلة للقيمة — ٦٧ر٠ ، وتكون نسبة النجاح ٢٥٪ مقابلة للقيمة + ٦٧ر٠ ، وهكذا . . وبهذا النوع من التدرج ، يمكننا أن نقارن صعوبة الفقرات التى حاولت الإجابة عنها جماعات مختلفة ، وهنا نتبع طريقة معينة فى تحديد الفرق فى المتوسط والانحراف المعيارى للقدرة بين الجماعتين . ولمعرفة المزيد عن هذه الطريقة يمكن الرجوع إلى :

Davis F.B., Item analysis Data, Harvard Education Papers No. 2, Graduate School of Education, Harvard University, Cambridge, Mass , 1946.

(ب) دلائل قدرة الفقرة على التمييز :

اقترح الكثيرون عدة دلائل توضح الدرجة التى تكون بها الفقرة فعالة فى التمييز بين أصحاب القدرة العالية وأصحاب القدرة المنخفضة ، ونحن نتعرف على مستويات القدرة هذه من الدرجة الكلية على الإختبار أو من محك خارجى آخر . وليس هذا مجال عرض هذه الدلائل جميعا إذ أنها إحصائية غالباً . ولكننا سنختار أبسطها وأميزها .

ونحن نهتم هنا بناحيتين : الأولى ، أننا قد نربط الأداء على الفقرة بأحد المقاييس المستمرة . وهذا المتغير المتصل أو المستمر غالباً هو الدرجة

على الإختبار ، الذى يشمل الفقرة المدروسة ، ككل . ولكن هذا المقياس المستمر قد يكون أحد المحكات الخارجية . وفى الناحية الثانية ، نربط الأداء على الفقرة بمتغير متقطع ، سواء أكان محكا أم كان الدرجة على الإختبار الكلى . والدرجة على الإختبار الكلى هى متغير مستمر ، ولكننا نستطيع تحويلها إلى متغير متقطع إذا حددنا درجة معينة من يفوقها يعتبر ذا قدرة عالية ومن يقل عنها يعتبر ذا قدرة منخفضة .

(ا) تحليل الفقرات فى حالة الدرجة المستمرة :

عند ما نحلل الفقرة على أساس متغير موزع إعتداليا ، تكون أمامنا عدة طرق يمكن استخدامها . وأول طريقة هنا ، نستخدم فيها درجة موزعة توزيعا مستمراً ، ولكنها طريقة شاقة . وفى الطريقة الثانية نحدد نقطة على المقياس المستمر تقسم الدرجات إلى مجموعتين عليا ودنيا ، وتعتبر المجموعتان متغيراً متقطعا . وفى الطريقة الثالثة نقارن بين مجموعتين متطرفتين ، والتطرف هنا يكون فيما يتعلق بالمتغير المستمر . وذلك بأن نحدد درجة مرتفعة جداً كل من يعلوها يعتبر متطرفا موجبا ، ثم نحدد درجة منخفضة جداً من يقل عنها يكون مجموعة متطرفة سالبة (تبعا لذيلى المنحنى الاعتدالى) ثم نقارن بين أداء أفراد هاتين المجموعتين على كل فقرة ندرسها .

وسنبدأ بالحديث عن المتغير المستمر كدرجة كمية Quantitative Score فعندما نرغب فى تحديد العلاقة بين كل فقرة وبين الدرجة الموزعة توزيعا مستمرا ، نستخدم معامل الإرتباط ذى الشعبتين (الثنائى Biserial أو الإرتباط الثنائى الأصيل Point Biserial) . وفى هذه الحالة نعتبر أداء الفقرة متغيرا متقطعا بمعنى أن الفرد يحصل فى الفقرة على درجة أو لا يحصل على شىء (صفر) وتكون الدرجة الكلية على الإختبار متغيراً مستمرا . ونوجد درجة الارتباط بين الفقرة المفردة والدرجة الكلية على الإختبار .

وإذا رأينا أن النجاح أو الفشل على الفقرة يمثل قدرة مستمرة ، وأن التقسيم إلى أفراد ناجحين وأفراد فاشلين فيما يتعلق بهذه القدرة هو تقسيم ممكن ، كان اختيارنا للنقطة التى تفصل بين المجموعتين متوقفا على صعوبة الفقرة المعينة . كان من المستحسن أن نستخدم معامل الإرتباط الثنائى . يحدث هذا فى حالة دلالة النجاح أو الفشل كعرض Symptom يخفى وراءه قدرة أو سمة معينة . ونحن نعلم أن قيمة معامل الإرتباط الثنائى لا تتوقف على نسبة المجموعة الناجحة ونسبة المجموعة الفاشلة وهذا يعنى أن دليل صدق الفقرة أو تميزها أمر منفصل تماما عن صعوبتها .

ومعامل الارتباط الثنائى الأصيل يتميز عن معامل الإرتباط الثنائى فى طبيعة الفروض التى يقوم عليها وفى تأثير صعوبة الفقرة على القيم الناتجة (وهى درجات على الإختبار) . فعامل الإرتباط الثنائى الحقيقى يفترض أن من ينجحوا فى الفقرة يكونوا جماعة متميزة عن جماعة من يفشلون . فهو يفترض أن الجماعتين تفترقان عن بعضهما إلا أن أعضاء كل جماعة يعتبرون متشابهين . وعلى هذا فقيمة الإرتباط الثنائى الأصيل فى تحليل الفقرات تعبر عن صعوبة الفقرة . ولكن افتراض أن المجموعتين متميزتان ليس صحيحاً دائماً . فإن الناجحين والفاشلين كلهم يمثلون مدى من القدرة تستطيع الفقرة إلى حد ما أن تقيسه . ويجب هنا أن نفرق بين صدق الفقرة وصعوبة الفقرة . واللبس بينهما يرجع إلى إستخدام معاملات مثل الإرتباط الثنائى الأصيل الذى يعطى قيمة تمثل الصعوبة والصدق معاً . ومن الأفضل أن نحصل على معلومات عن الصدق أو القدرة التميزية على حدة وعن صعوبة الفقرة على حدة ثم نجمع بين هذه الحقائق بطريقة واضحة . وأخيراً فإن حساب معاملات الثبات الثنائية بالنسبة لكل فقرة عمل شاق جداً .

٢ ــ تحليل الفقرات فى حالة الجماعات المتقطعة :

فى عمليات تحليل الفقرات نضحى بالطبيعة المستمرة المتغير الذى نربط به الفقرة فى سبيل تسهيل عمليات الحساب . فنعتبر أن من يحصلون على درجات عالية يمثلون مجموعة منفصلة عن مجموعة من حصلوا على درجات منخفضة برغم أن متغير الدرجة قد يكون مستمراً غالباً . ويحسن أن تقسيم الدرجة التى يختارها الأفراد إلى نصفين وتستخدم نفس الخطوات عندما يكون المتغير متقطعاً بالفعل مثل المحك (كالتقسيم إلى مريض وسوى أو ناجح وفاشل) :

فإذا قسمنا المجموعة بهذه الطريقة ، كانت أمامنا عدة طرق لتقييم الفقرة . وأول هذه الطرق بسيط جدا ويخدم كثيرا من أغراض واضعى الاختبارات غير الرسمية informal Tests ويستعين بها المحللون فى تحسين وتعديل الإختبارات . وفى هذه الطريقة نقوم بحساب النسبة المئوية الناجحة على الفقرة من أفراد المجموعة التى حصلت على أعلى الدرجات فى الإختبار ككل ثم نحسب النسبة المئوية للأفراد الناجحين فى نفس الفقرة من المجموعة التى حصلت على أدنى الدرجات فى الإختبار . ثم نعقد مقارنة بين هاتين النسبتين المئويتين . وهذه الطريقة مفيدة فى استبعاد الفقرات غير الصالحة . وخطوات هذه العملية هى : عزل أوراق إجابة المفحوصين الذين حصلوا على درجات عالية جداً عن أوراق إجابة من حصلوا على درجات منخفضة جداً . ثم جمع عـــدد الإجابات الصحيحة ، وغير الصحيحة والفقرات المتروكة لصعوبتها ، والفقرات التى لم يستطع الفرد الإجابة عنها لضيق الوقت . وسواء أكنا نستخدم النسبة المئوية الخام للإجابات الصحيحة ، أو النسبة المئوية المعدلة لاستبعاد تدخل الصدفة ، فإننا نقارن بين المجموعتين فى الإجابة عن كل فقرة على هذه الأسس .

وقد أجرى تعديل على هذه الطريقة بتقسيم الأفراد الذين طبق عليهم

الإختبار إلى مجموعة ناجحة وأخرى فاشلة ، ثم تحديد نسبة الناجحين فى الإجابة عن كل فقرة ضمن المجموعتين الكبيرتين كل على حدة . ثم يستخدم عامل الارتباط الرباعى أو معامل phy لمعالجة هذه النسب المئوية . ويجب أن نضع فى ذهننا أننا ندرس استجابة فقرة وليس أننا ندرس متغيراً متقطعاً منفصلا . وعلى هذا تكون لدينا أربعة نسب مئوية توضع فى جدول من أربعة خلايا ويحسب الارتباط بين هذه الخلايا إما بمعامل الارتباط الرباعى أو بمعامل فاى . وإختبار أحد معاملى الارتباط متوقف على افتراض استمرار توزيعى المتغيرين موضع الدراسة . فإذا افترضنا أنهما مستمران وموزعان اعتدالياً ، استخدمنا الارتباط الرباعى . وإذا افترضنا أن المتغيرين متقطعان استخدمنا ارتباط فاى . ومعامل الارتباط الرباعى لا يتأثر بصعوبة الفقرة ، بينما يمثل ارتباط فاى ، كما كان الحال فى الارتباط الثنائى الأصيل ، مستوى صعوبة الفقرة المعينة .

(ج) تحليل الفقرات فى حالة الجماعات المتطرفة :

إذا قررنا أن نترجم درجاتنا المستمرة إلى فئتين نقسم المجموعة الكبيرة تبعاً لهما إلى مجموعتين عليا ودنيا ، كنا بصدد إمكان تحليل هاتين المجموعتين المتطرفتين ، وإسقاط باقى الأفراد فى المنطقة المتوسطة . فإذا كانت العلاقة بين درجة الاختبار ودرجة الفقرة علاقة مستقيمة Linear (أى أن زيادة أحدهما تؤدى دائماً إما إلى نقص الأخرى وإما إلى زيادتها) بحيث أن النسبة المئوية للناجحين فى الفقرة تتزايد بمعدل ثابت تتزايد به الدرجة الكلية ، كان أخذ مجموعتين من الطرفين مؤدياً إلى تضخيم الفروق التى نلاحظها فى الفقرة المفردة . فالفقرة تستطيع التمييز الواضح بين العشر الأعلى Top tenth والعشر الأدنى bottom tenth فى المجموع السكانى المدروس أكثر مما تستطيع بين النصف الأعلى والنصف الأدنى . وبرغم أن لهذه الطريقة ميزة توضيح

الفروق وإبرازها إلا أنها تتجاهل الأفراد فى الأقسام الثمانية الأخرى (المحصورة بين العشر الأعلى والعشر الأدنى) فبالرغم من أن الفقرة تميز بوضوح بين أعلى ١٠ أفراد وأدنى ١٠ أفراد فى مجموعة من ١٠٠٠ فرد أكثر مما تميز بين أعلى ٥٠٠ وأدنى ٥٠٠ ، إلا أننا نتوقع أن مقارنة الفقرات المختلفة إذا اعتمدت على العُشرين لا يمكن الاعتماد عليها على عكس المقارنة بين النصفين . وعلى ذلك فإن علينا أن نبحث عن التميز مع الإبقاء على الثبات . فهل هناك نقطة فاصلة يكون فيها الاتزان أفضل بين حدة التميز وبين ثبات القيم التى نحصل عليها ؟ بينت كيلى فى :

Kelley T. L., "The Sellection of the Upper and Lower Groups for the Validation of Test Items", J. Educ. Psychol., 30, 17—24 (1939).

إن النسبة بين الفروق والخطأ المعيارى تبلغ أقصاها عندما تتضمن المجموعة العليا ٢٧ ٪ والمجموعة الدنيا ٢٧٪ تقريباً من كل المجموع السكانى المدروس . وأن زيادة أو قلة النسبة المئوية عن هذا الحد تؤدى إلى إنخفاض الدقة التى نرتب بها الفقرات مبتدئين من أقل الفقرات قدرة على التميز إلى أكثرها . وهكذا ندرس ٢٧ + ٢٧ = ٥٤ ٪ من الحالات المتطرفة ونرفض ٤٦ ٪ من الحالات المتوسطة ، ونسهل عمليات الحساب ونزيد دقة نتائجنا .

(هـ) تحليل الفقرات فى حالة محك متقطع :

قد يتطلب الحال أن نربط النجاح على الفقرة بمحك منقطع مثل النجاح فى الدراسة أو العمل (بالتقسيم إلى ناجح وفاشل مثلا) . وهنا لا تقسم المجموعة المختبرة ، غالباً إلى قسمين متساويين ، ونوجد معامل الارتباط إما باستخدام الارتباط الرباعى أو بارتباط فاى .

الصـــدق المتبادل

يشير الصدق المتبادل Cross Validation إلى حقيقة أن صدق الاختبار يجب أن يتحدد بعينة من الأشخاص تختلف عن العينة التى جربت عليها فقراته عند اختيارها . فأى معامل صدق حسب على نفس العينة التى استخدمت فى أغراض اختيار الفقرة يتضخم على حساب أخطاء الصدفة فى حدود هذه العينة ويرتفع بشكل مبالغ فيه . والحق أن معامل الصدق المرتفع للاختبار فى مثل هذه الحالات ، يخفى وراءه أن الاختبار ليس صادقاً إطلاقاً فى التنبؤ بمحك معين .

فإذا افترضنا أن لدينا ١٠٠ طالب طب ، اخترنا منهم أعلى ٢٧ وأدنى ٢٧ فى درجاتهم الدراسية ، ليمثلوا محكاً من مجموعتين متقابلتين . فإذا قارنا بين هاتين المجموعتين فى عدد من السمات التى لا تتعلق فعلا بالنجاح فى كلية الطب ، فسنجد بالتأكيد فروق الصدفة . فقد نجد مثلا زيادة عدد الحضريين عن الريفيين داخل مجموعة المحك العليا . فإذا أعطينا درجة عن الإقامة أو الميلاد فى إقليم حضرى زادت درجات مجموعة المحك العليا عن مجموعة المحك الدنيا أكثر من ذى قبل . وليس فى هذا دليل على صدق الاختبار كأداة للتنبؤ . فقد اخترنا عوامل تتغير بالصدفة وتميز هـــذه العينة بالذات ، بمعنى أننا لو اخترنا مجموعة أخرى فقد نجد أن الريفيين فيها أكثر .

وترفع معاملات الصدق ، بهذا الشكل المبالغ فيه ، فى حالة اختبارات الشخصية حيث لا تصحح الدرجة أساسا على أنها صحيحة أو خاطئة . وفى هذه الحالة ، يتحدد اختيار وتصحيح الفقرات باستجابات العينة الأولى . وعلى هذا تحصل الاستجابات فى مجموعة المحك العليا على الوزن + ١ ، وتشيع الإجابات التى تحصل على الوزن — ١ فى المجموعة الدنيا . وفى مثل

هذه الأحوال ، تتدخل فروق الصدفة فى أى من جماعتى المحك ، وتكون العلاقة بين الاختبار والمحك عالية بدرجة مبالغ فيها . وفى :

Kurtz, A. K. A Research test of the Rorchach test. Personnel Psychol. 1948.

نجد مثالا يدلنا على الحاجة إلى بحث الصـــدق المتبادل . فقد قام Kurty فى بحث على اختبار الرورشاخ لتحديد ما إذا كان هذا الاختبار يفيد فى انتقاء مديرى المبيعات فى إحدى شركات التأمين ، فطبق الاختبار على ٨٠ مديرا . وقد اختير هؤلاء المديرين من بين عدة مئات من موظفى ٨ شركات تأمين . وحددت جماعة المحك العليا بحيث كانت ٤٢ مديرا ناجحين فى مقابل جماعة المحك الدنيا التى تكونت من ٣٨ مديرا اعتبروا فاشلين . وقام أحد الخبراء فى اختبار الرورشاخ بدراسة السجلات الثمانين ، وقد اختار مجموعة من العلامات تبلغ ٣٢ علامة هى مميزات فى الاستجابة تكررت بنسبة أكبر فى إحدى جماعتى المحك أكثر مما تكررت فى الأخرى . والعلامات التى تكررت كثيراً فى جماعة المحك العليا أعطيت الوزن + ١ عند من استجاب بها والوزن صفر إذا لم يستجب بها عضو الجماعة العليا ، والإستجابات التى تكررت أكثر فى جماعة المحك الدنيا أعطيت الوزن — ١ إذا حدثت والوزن صفر إذا لم تحدث فى استجابات عضو هذه الجماعة . وقد بلغ عدد العلامات التى تكررت فى المجموعة العليا ١٦ علامة ووزنها + ١٦ كحد أقصى ، كما بلغ عدد العلامات التى تكررت فى المجموعة الدنيا ١٦ علامة ومجموع وزنها كحد أدنى — ١٦ . أى أن الدرجة الكلية انحصرت نظريا فى المدى بين — ١٦ ، + ١٦ .

وعند إعادة تطبيق مفتاح التصحيح المبنى على ٣٢ علامة على المجموعة الأصلية وعددها ٨٠ شخصا وجدنا أن ٧٩ منهم قد وزعوا نفس التوزيع السابق إلى مجموعة عليا وأخرى دنيا . وعلى هذا كان معامل الارتباط بين

بين درجة الاختبار والمحك قريباً جداً من ١ . وعند ما بحث الصدق المتبادل على عينة مقارنة ثانية مكونة من ٤١ مديرا ، ٢١ فى المجموعة العليا و٢٠ فى المجموعة الدنيا فانخفض معامل الصدق إلى ٢ ٠ر٠ . ومن هنا اتضح أن المفتاح الذى أنشئ فى العينـــة الأولى كان غير صادق فى اختيار هؤلاء المديرين .

أوضح البحث السابق مدى زيف النتائج متأثرة بعوامل الصدفة . وها كم بحثا آخرَ يؤيد هذا نجده فى :

Cureton, E. E. Validity. Reliability and baloney. Educ, Psychol. Measnet 1950.

والمفحوصون فى هذه الدراسة ٢٩ تلميذاً بالسنة النهائية فى التعليم الثانوى وبالجامعة يدرسون المقاييس النفسية . والمحك الذى طلب التنبؤ به كان متوسط درجة التلميذ . والفقرات تكونت من ٨٥ قصاصة من الورق مرقمة من ١ حتى ٨٥ وكل الأرقام على وجه واحد ولإعطاء كل مفحوص درجة ، كانت القصاصات توضع فى وعاء وتهز ثم تسقط على المائدة . وكل القصاصات التى سقطت بحيث ظهر الوجه المرقم اعتبرت دالة على وجود تلك الفقرة المعينة فى أداء التلميذ للاختبار . أدى إلقاء القصاصات (وعددها ٨٥) مرات عددها ٢٩ إلى سجلات كاملة لكل تلميذ ، يبين وجود أو غياب كل فقرة أو علامة فى الاستجابة . هذه العملية ، بالطبع ، هى مجرد أداة سريعة للحصول على نتائج بمحض الصدفة . وقد أقيم تحليل لهذه الفقرات باعتبار متوسط الدرجة الحقيقية لكل تلميذ محكا . وعلى هذا الأساس اختار الباحث ٢٤ فقرة من ٨٥ فقرة فى الاختبار ، ومن هذه الفقرات الـ ٢٤ وجدنا ٩ فقرات تكررت أكثر بين التلاميذ الذين حصلوا على أعلى الدرجات ، و ١٥ فقرة شاعت فيمن حصلوا على أدنى الدرجات : وأعطى الوزن + ١ لكل فقرة من الفقرات التسعة ، وكان وزن كل

فقرة من الـ ١٥ فقرة هو – ٠١ مجموع أوزان هذه الفقرات بالنسبة لكل تلميذ يكون درجته الكلية . وبالرغم من معرفة أن الصدفة حرفت درجات هذا الاختبار إلا أن معامل ارتباطها بمحك الدرجة الحقيقية للتلاميذ فى المجموعة الأصلية وعددها ٢٩ بلغ ٠ر٨٢ . ومثل هذه النتائج تتفق مع بحث اختبار الرورشاخ . ففى كلتا الحالتين ، نشأ الارتباط الظاهر بين درجة الاختبار والمحك من استخدام نفس فروق الصدفة سواء عند تصميم مفتاح التصحيح أو عند مراجعة الصدق .

ويجب أن نلاحظ أن هذا الارتفاع الكبير لمعامل الصدق لا يحدث وحسب فى اختيار الفقرات بل وأيضاً فى اختيار الاختبارات لتكوين بطارية . فإن معامل الارتباط المتعـــدد بين البطارية والمحك ينخفض عادة عندما يعاد حسابه على عينة جديدة . وهذا الانخفاض المفاجئ سيكون صـــغيراً فى حالة إنشاء البطارية ، ما دامت نسبة الاختبارات المحذوفة فى العملية التالية صغيرة نسبياً ، وفى تحليل الفقرات ، من المعتاد أن نستبعد فقرات أكثر من تلك التى نبقى عليها . وعندما يكون عدد الفقرات الأصلية كبيراً ونسبة الفقرات التى أبقيناها صغيرة ، تزيد فرصة تضخم قيمة فروق الصدفة وبالتالى نحصل على معامل صدق مرتفع بشكل زائف . ولهذا السبب تزيد الحاجة إلى الصدق المتبادل عند تحليل الفقرات منها عند إنشاء البطارية .

وفى أى من الحالتين يزيد الزيف كلما صغرت العينة . فما دام معامل الصدق راجعاً إلى تجمع أخطاء العينة ، فإن زيادة هذه الأخطاء تؤدى إلى زيادة هذا الزيف ، ويحدث هذا بالطبع كلما صغر حجم العينة ، فعندما تتكون العينة الأصلية من عدة آلاف حالة ، قد يبلغ معامل الصدق بطريقة الصدق المتبادل رقماً صغيراً جداً . وعلى هذا فإن العوامل المحددة الأساسية هى حجم العينة ، وعدد الفقرات ، ونسبة الفقرات التى أبقيت وتصبح

الحاجة ماسة عندما تكون العينات صغيرة ، والفقرات الأصلية كثيرة ، ونسبة الفقرات التى أبقيناها قليلة . وفى مثل هذه الظروف ، نجد معامل الصدق المحسوب من نفس العينة المستخدمة فى اختيار الفقرة معاملا كاذباً .

ثبات الفقرات :

الاختبار مقياس وحداته الفقرات يعتمد صدقه على صدقها اعتماد ثباته على ثباتها . وعلى ذلك فكما بحثنا صدق الفقرة ، باعتباره مؤدياً إلى صدق الاختبار ، نبحث ثبات الفقرة بعدة طرق أهمها طريقة إعادة الاختبار وطريقة الاحتمال المنوالى Modal Probability .

وفى طريقة إعادة الاختبار نسجل نتائج الإجراء الأول فى جدول كل فرد فى عمود وكل فقرة فى صف ويقسم الجدول إلى قسم للإجابات الصحيحة وآخر للإجابات الخاطئة . وترصد نتائج إعادة الإختبار بنفس الطريقة ثم يحسب معامل الارتباط الرباعى بين مرتى الإجراء .

وفى طريقة الاحتمال المنوالى ، نستخدم المعادلة التالية :

$$\text{معامل ثبات الفقرة} = \frac{\text{ه}}{\text{ه} - ١} \left(\text{ل} - \frac{١}{\text{ه}} \right)$$

حيث ه هى عدد متغيرات الإجابة التى يستطيع أو يجب أن يختار إحداها المفحوص بالنسبة للفقرة وليكن (٥) .

و ل هى الاحتمال المنوالى أى النسبة بين أكبر عدد اختار إحدى متغيرات الإجابة وبين مجموع عدد الأفراد الذين طبقت عليهم هذه الفقرة وليكن عدد من اختاروا إحدى متغيرات الإجابة هو (٢٥) وعدد من طبقت عليهم هذه الفقرة هو (٥٠) فتكون النسبة ل = $\frac{٢٥}{٥٠}$ = ٥٠ر٠

إذن معامل ثبات هذه الفقرة = $\frac{٥}{٥ - ١}$ (٥٠ر٠ − $\frac{١}{٥}$)

= $\frac{٥}{٤}$ (٥٠ر٠ − ٢٠ر٠) = $\frac{٥}{٤}$ (٣٠ر٠) = $\frac{٣}{٨}$

= ٣٧٥ر٠ = ٣٨ر٠ تقريباً .

وتفيد طريقة الاحتمال المتوالى كما رأينا فى حالة إيجاد ثبات الفقرات التى تتطلب الإجابة عنها الإختيار من عدة إجابات أو الإجابة بإحدى الاحتمالات التى يستطيع الباحث أن يحصرها . ولمعرفة المزيد عن هذه الطرق ارجع إلى :

Hollzinger, K. J. Reliability of a single test item, J, Ed. Vol 23. No 6.

وزن الإجابات عن الفقرات :

المشاكل التى صادفتنا عند وزن الإختبارات التى تتكون منها البطارية تعود لتواجهنا عند وزن الفقرات التى يتكون منها الاختبار . وتبعاً لإحدى المعادلات الإحصائية ، وهى معادلة الانحدار المتعدد **Multiple Regression Equation** ، يجب علينا أن نربط الفقرات بعضها بالبعض الآخر كما نربط كل فقرة منها بالمحك ، ولكن هذا الأمر جد مرهق فهو يتطلب القيام ، عادة ، بحساب آلاف معاملات الارتباط . وللتغلب على هذه الصعوبة اشتقت عدة معادلات . ولكننا دائماً نراعى أن يرتفع معامل الارتباط بين كل فقرة وبين المحك إلى أقصى حد ، فى حين تنخفض معاملات الارتباط الداخلية بين

الفقرات وبعضها ، وإلا فإن الفقرة التى ترتبط ارتباطاً منخفضاً بالمحك يجب أن ترتبط ارتباطاً عالياً بفقرة أخرى أو بالفقرات الأخرى حتى تصبح ذات قيمة فى التنبؤ بالمحك .

ولتبسيط معادلة الانحدار قامت محاولات كثيرة أهمها الطريقة التى نجدها فى :

Richardson, M. W. and Adkins, D. C. : A Rapid Method of selecting test items J, educ. Psychol. 1938, 29.

وهذه الطريقة تحسب وزن كل فقرة بتطبيق المعادلة الآتية عليها :

$$و = \frac{ر_{ص.ع} - ر_{س.ع}\, ر_{س.ص}}{\sqrt{ا\, ب\,(ر_{س.ص} - ر_{ص.ع}\, ر_{س.ع})}}$$

حيث (و) هو وزن الفقرة .

، ($ر_{ص.ع}$) هو معامل الارتباط بين الفقرة والمحك وليكن (٩٠ر٠)

، ($ر_{س.ع}$) هو معامل الارتباط بين الفقرة ودرجة الاختبار الكلى الشامل لهذه الفقرة وليكن (٣٠ر٠)

، ($ر_{س.ص}$) هو معامل الارتباط بين درجة الاختبار الكلى والمحك ولتكن (٦٠ر٠)

، (ا) هى نسبة الناجحين فى الفقرة ولتكن (٨٠ر٠)

، (ب) هى نسبة الفاشلين فى الفقرة ولتكن (٢٠ر٠)

$$و = \frac{٩٠ر٠ - ٣٠ر٠ \times ٦٠ر٠}{\sqrt{٨٠ر٠ \times ٢٠ر٠\,(٦٠ر٠ - ٩٠ر٠ \times ٣٠ر٠)}}$$

$$= \frac{٩٠ر٠ - ١٨ر٠}{\sqrt{١٦ر٠}\ (٦٠ر٠ - ٢٧ر٠)}$$

$$= \frac{٧٢ر٠}{٤ر٠\ (٣٣ر٠)}$$

$$= \frac{٧٢ر٠}{١٣٢ر٠}$$

$$= \frac{٦}{١١}$$

$$= ٥٦ر٠$$

إذن وزن الفقرة هو ٥٦ر٠ ؛

أى أن من يجيب عن هذه الفقرة إجابة صحيحة يحتمل له النجاح فى العمل أو الدراسة التى تطبق الاختبار لانتقاء الملتحقين بها ويكون الاحتمال بمقدار ٥٦ ٪ واحتمال فشل نفس الفرد هو ٤٤ ٪ كما أن هذا يعنى أنه لو أجاب ١٠٠ فردا إجابة صحيحة عن هذه الفقرة توقعنا أن ينجح منهم (فى المحك الذى نتنبأ له) ٥٦ فرداً وأن يفشل ٤٤ فرداً ؛

وتطبق هذه الطريقة فى حالة الاهتمام بمعادلات الارتباط الداخلية بين الفقرات ، أما إذا أهملنا هذه المعاملات ، أو افترضنا أنها متسارية بين الفقرات المختلفة ، فإنه يمكننا أن نعطى وزناً للتصحيح بطريقة جيلفورد Guilford والمعادلة هى :

$$\text{و} = \frac{\text{نسبة جماعة المحك العليا} - \text{نسبة جماعة المحك الدنيا}}{\text{النسبة من الجماعة العليا} \times \text{النسبة من الدنيا}} + ٤$$

حيث (و) هى وزن طريقة معينة فى الإجابة عن الفقرة .

ونسبة الجماعة العليا ، فى المحك أو التى حصلت على درجات عالية فى الاختبار ككل ، هى النسبة المئوية من الناجحين فى المحك التى أجابت بهذه الطريقة عن تلك الفقرة ولتكن (٧٠ر٠) أى أن (٣٠ر٠) من الناجحين أجابوا إجابات مختلفة .

ونسبة جماعة المحك الدنيا هى نسبة الأفراد من جماعة الفاشلين ، فى المحك أو الاختبار ككل ، التى أجابت بالطريقة نفسها عن الفقرة نفسها ولتكن (٤٠ر٠) أى أن (٦٠ر٠) من الفاشلين أجابوا بطريقة مختلفة .

$$\therefore \text{و} = \frac{٧٠ر٠ - ٣٠ر٠}{٧٠ر٠ \times ٣٠ر٠} + ٤$$

$$= \frac{٤ر٠}{٢١ر٠} + ٤$$

$$= ٩ر٥$$

ونتيجة هذه المعادلة تتراوح بين صفر ، ٨ ؛ وإذا كان الوزن الناتج هو ٤ كان ارتباط الفقرة بالمحك صفراً . وهذه المعادلة تطبق بوجه خاص فى حالة تساوى مجموعتى المحك العليا والدنيا فى عدد الأفراد فى كل منهما . وإذا لم تتساوى المجموعتان أو كانت نسبة الناجحين فى المحك الذين أجابوا إجابة معينة على فقرة معينة مساوية لنسبة الفاشلين الذين أجابوا بهذه

الطريقة ، فى هذه الحالة سيكون الوزن ٤ والارتباط بين الفقرة والمحك صفراً والخطأ المعيارى كبيراً . أى فى حالة فرض الصفر أو العدم null hypothesis وهنا يحسن استخدام المعادلة التالية :

$$ع_و = \frac{٢}{\sqrt{ن\ ا\ ب}}$$

حيث $ع_و$ هى الوزن المعدل للفقرة

، ن هى مجموع عدد الفاشلين وليكن (٢٠) والناجحين وليكن (٣٠)

، ا هى نسبة الناجحين فى المجموعة ولتكن (٢٠ر٠) التى أجابت بطريقة معينة :

، ب هى نسبة الفاشلين فى المجموعة ولتكن (٣٠ر٠) التى أجابت بنفس الطريقة .

$$\therefore \text{الوزن المعدل} = \frac{٢}{\sqrt{٥٠ \times ٢٠ر٠ \times ٣٠ر٠}}$$

$$= \frac{٢}{\sqrt{٣}}$$

$$= \frac{٢}{١ر٧٣٢١}$$

$$= ١ر١٥$$

أهمية وزن الإجابات :

وتنحصر أهمية وزن الاستجابات عن الفقرات فى أن اختبارات الشخصية عامة قد اتبعت طرقاً عقيمة فى وزن الاستجابات على فقراتها . إلا أن بعض الاختبارات قد أفاد من وزن الإجابات ، مثل اختبار سترونج للميول المهنية Strong, Vocational Intrests Blank (V.I.B.) واختبار برنروتير للشخصية Bernreuter Personality Test ، فى رفع ثباتها . والتصحيح غير الموزون Unweighted Scoring هو إعطاء الاستجابة صفراً أو (١) . كما استفادت الاختبارات من وزن استجاباتها فى رفع صدقها . وعلى هذا ، يكون من المهم أن نتأكد من أن الفقرات ترتبط ارتباطا جوهريا بالمحكات : وبهذه الطريقة يمكننا ، بتطبيق الاختبار على عينات كبيرة ، أن نختار فقرات يمكن الاعتماد عليها وأن ننشأ اختبارا يمكن الاعتماد عليه مادمنا قد طبقناه على عينات صدق ممثلة أخرى غير العينات التى استخدمناها فى اختيار الفقرات .

استخدام صدق الفقرات

قلنا إنه عند تحليل الفقرات نهتم بأمرين ، أولهما تحديد مستوى صعوبة الفقرات والعمليات المستخدمة فى حسابه هى عمليات الاتساق الداخلى . والأمر الثانى هو القدرة التى بها تميز الفقرة بين المفحوصين ، تميزا يتفق مع تميزهم فى محك آخر يقيس ما تقيسه الفقرة وهو غالباً محك خارجى آخر كالنجاح فى العمل أو الدراسة ، هذه العملية هى إيجاد معامل صدق الفقرة .

وكلما زاد عدد الفقرات وكانت غير متجانسة ، أى انخفضت معاملات الارتباط الداخلية بينها ، زاد احتمال أنها تقيس جوانب مختلفة من القدرة . وهنا لا يفيدنا إيجاد الارتباطات الداخلية بغرض تحديد الاتساق الداخلى ،

بل نهتم بإيجاد صدق كل فقرة . فنعامل كل فقرة بإيجاد معامل الارتباط بينها وبين المحك الخارجى . وهكذا يكون لدينا معامل صدق لكل فقرة وكأن كل فقرة أصبحت اختباراً منفصلا . وأصبح لدينا إذن عدد كبير من معاملات صدق اختبارات عديدة مختصرة جدا . وتكون المشكلة ، إذن ، هى جمع هذه الاختبارات فى اختبار واحد كبير على أن يتمتع بأعلى معامل ارتباط مع المحك . ونحن هنا نقوم بنفس العمل الذى قمنا به عند جمع عدة اختبارات فى بطارية واحدة فنراعى ارتفاع معامل الصدق وانخفاض درجة توحد uniquenes الفقرات فى قياس الجانب المعين . بحيث تصبح الفقرة صادقة إلى درجة كبيرة وتقيس شيئاً يختلف عن الذى تقيسه فقرة أخرى .

فيعد الاختبار (فى حالة البطارية) أو تعد الفقرة (فى حالة الاختبار) ذات قيمة تدفعنا إلى فقرات أخرى أو اختبارات أخرى إذا توفر فيها الشرطان الآتيان : (١) ارتفاع معامل صدقها (٢) انخفاض معامل ارتباطها مع الفقرات الأخرى . فتزيد قيمة الفقرة كلما كانت تقيس جانباً لا تقيسه فقرات أخرى . فإذا أعطينا للفقرات أوزاناً وكانت متساوية ظهرت مشكلة اختيار أى هذه الفقرات ضمن الاختبار . على أن يكون العدد المختار من الفقرات اختباراً يحصل فيه الأفراد على درجات ترتبط أعلى ارتباط مع المحك أى تبلغ أقصى درجة ممكنة من الصدق . وحل هذه المشكلة يمكن أن يتم بالطرق الثلاثة الآتية :

الطريقة الأولى :

يمكننا أن نتجاهل معاملات الارتباط الداخلية بين الفقرات تماماً . وهكذا يكون الأساس الوحيد فى اختيار الفقرات هو صدق كل منها على حدة . وعلى ذلك يجب أن يحدد مستوى مناسباً للصدق إذا بلغته الفقرة

كانت صالحة وإلا استبعدت من الاختبار . ويحسن ، فى أنواع معينة من المقاييس كالاستفتاءات وتواريخ الحياة والاستبيانات ، أن تحلل كل إجابة محتملة عن الفقرة بالطريقة الآتية :

لنفرض أنه فى أحد اختبارات التشخيص **diagnosis** وجدنا الفقرة التالية :

ما هو حجم الجماعة التى تميل إلى الانضمام إليها ؟

(أ) ٢ شخصاً

(ب) من ٢ — ٥ شخصاً

(ج) من ٥ — ٨ شخصاً

(د) من ٨ — ١٢ شخصاً

(هـ) أكثر من ١٢ شخصاً

فعلينا هنا أن نحلل هذه الإجابات المحتملة للفقرة الواحدة فإذا فرضنا أن ٥٠ ٪ من الأفراد الذين أجابوا مختارين الإختيار (أ) كانوا مصابين باضطراب معين ، وأن ٤٤ ٪ ممن اختاروا (ب) كانوا مرضى بأحد الأمراض . وأن ٣٢ ٪ ممن اختاروا (ج) كانوا مضطربين ، وأن ٢٣ ٪ من الأفراد الذين وافقوا على الاختيار (د) كانوا يعانون من أزمات نفسية . وأن ١٥ ٪ ممن أجابوا مختارين (هـ) كانوا مصابين بأحد النوبات . كل هذه الإجابات إذن ترتبط بالمحك الخارجى وهو هنا التشخيص . أى أن كل إختيار له درجة من الصدق ويعطى وزناً يحسب هذا الصدق . وتكون هذه الفقرة مناسبة لاختبار غرضه التشخيص .

فإذا كانت لدينا فقرات ذات معاملات صدق مختلفة كيف نفاضل بينها ؟ وما هو معامل الصدق الذى نتوقع أن يصل إليه الاختبار المكون من هذه الفقرات ؟ لنفرض أن لدينا ٥ فقرات (أ) معامل صدقها ٢٥ر٠ ،

(ب) معامل صدقها (٢٠ر٠) ، (جـ) معامل صدقها ١٥ر٠ ، (د) معامل صدقها ١٠ر٠ ، (هـ) معامل صدقها ٠٥ر٠ : فباستخدام إحدى المعادلات الإحصائية (ارجع إلى كتب الإحصاء) يمكننا أن نحسب معامل صدق الاختبار الذى يشمل الفقرات (أ) ، (ب) والاختبار الذى يتكون من (أ) ، (ب) ، (جـ) ، والاختبار الذى يتكون من (أ) ، (ب) ، (جـ) ، (د) . والاختبار الذى يتكون من (أ) ، (ب) ، (جـ) ، (د) ، هـ : ذلك بفرض أن معاملات الارتباط بين كل فقرة والأخرى من هذه الفقرات الخمس هو ٥٠ر٠ دائماً مرة و ٢٠ر٠ دائماً فى أخرى وصفر من مائة فى الثالثة . بهذه نحصل على الجدول التالى :

الفقرات المتضمنة	معامل الإرتباط بين كل الفقرات		
	٥٠ ر٠	٢٠ر٠	٠٠ر٠
(ا)	٢٥٠ر٠	٢٥٠ر٠	٢٥٠ر٠
(ا) ، (ب)	٢٦٠ر٠	٢٩٠ر٠	٣١٨ر٠
(ا) ، (ب) ، (حـ)	٢٤٥ر٠	٢٩٣ر٠	٣٤٦ر٠
(ا)،(ب)،(حـ)،(د)	٢٢٢ر٠	٢٧٧ر٠	٣٥٠ر٠
(ا)،(ب)،(حـ)،(د)،(هـ)	١٩٤ر٠	٢٥٠ر٠	٣٣٥ر٠

جدول (٥) يوضح صدق الدرجة الكلية عن خمسة فقرات لكل معامل صدقها . عن

Thorndike, R. L., Personnel sellctionl N. y. : John wiley sons 1949..

ومن الجدول يتضح لنا :

١ ــ أنه كلما انخفض معامل الارتباط الداخلى بين الفقرات ، على أن تكون كل منها صادقة ، زاد معامل صدق الاختبار ككل .

٢ ــ أن زيادة الفقرات غير الصادقة كالفقرة (ﻫ) يخفض معامل صدق الاختبار ككل .

٣ ــ أن زيادة فقرة ذات معامل صدق صغير لا يزيد من صدق الاختبار بل يخفضه غالباً كما فى إضافة الفقرة (د) .

٤ ــ أنه كلما زاد طول الاختبار زاد صدقه ، على أن تنخفض معاملات الارتباط الداخلية بين فقراته .

فالمشكلة إذن هى اختيار مجموعة الفقرات التى تكون اختباراً يحصل فيه أفراد العينة ، التى سيطبق عليها ، على درجات تتمتع بأعلى معامل صدق ممكن . وهذه المشكلة ، بدورها ، تتضمن مشكلة أخرى وهى تحديد القيمة المناسبة التى يجب أن يبلغها صدق الفقرة حتى يمكن إضافتها إلى عينة القدرة المقاسة أى إلى الفقرات الأخرى فى الاختبار . وهنا يجب أن ندخل فى الحساب انحدار معاملات صدق الفقرات من العينة الأولى (سواء كانت العينة هى فقرات الاختبار الأصلى الذى أخذت منه لتضاف إلى عينة جديدة هى فقرات الاختبار الذى نود إطالته أم كانت العينة هى جماعات المفحوصين التى كان الاختبار يطبق عليها ونريد تطبيقه على عينات أخرى كما يحدث عند نقل اختبار أجنبى إلى بيئة أخرى)،ولا يتسع المقام لشرح هذه العمليات الإحصائية التى يمكن الرجوع فيها إلى كتب الإحصاء .

وإذا كانت لدينا مجموعة من الفقرات لا نعرف معاملات الارتباط الداخلية بينها ، علماً بأننا نسلم بأثر الفقرات على صدق الدرجة الكلية على الاختبار ، فلا يمكننا حل هذه المشكلة بطريقة تحليلية إذ أننا لا نستطيع بهذه الطريقة تحديد مدى هذا الأثر .

الطريقة الثانية :

يمكننا معالجة مشكلة انتقاء الفقرات لوزنها فى مفتاح التصحيح بأن نضع فى الاعتبار معاملات الارتباط الداخلية بين الفقرات بعد عمل سلسلة من

عمليات التقريب . والعملية العامة فى هذا هى وضع مفتاح أولى للاختبار يستند إلى مجموعة صغيرة من الفقرات الأكثر صدقاً . ويحصل كل فرد على درجة باستخدام مفتاح التصحيح هذا . ويحدد صدق هذه الدرجات . ثم يحسب ارتباط كل فقرة بهذه الدرجة ، فباستخدام المفتاح الجديد لإعطاء الدرجات نكون قد استخدمنا الدرجة الكلية كمتغير متصل يفيد كمحك فى تحليل الفقرات . وقد وضح Flangan هذه الطريقة . وهكذا نحصل بالنسبة لكل فقرة على :

١ ــ معامل ارتباطها بالمحك .

٢ ــ معامل ارتباطها بالمجموعة الأولى من الفقرات الموزونة .

٣ ــ ونعرف أيضاً صدق الدرجة عن المجموعة الأولى من الفقرات .

٤ ــ وبهذه الطريقة نستطيع أن نحسب بالنسبة لكل فقرة معامل ارتباطها الجزئى بالمحك عندما تصبح الدرجة عن الفقرات التى اخترناها أولا مقداراً ثابتاً Constant .

٥ ــ وهذا يدلنا على فائدة كل فقرة بالنسبة لصدق المجموعة الأولى من الفقرات . أى أن تحليل كل من الفقرات التى اخترناها أولا يدلنا على ما تضيفه الفقرة إلى باقى الفقرات من رفع معامل صدقها .

ونستطيع أن نختار الآن مجموعة ثانية من الفقرات لنصححها ، مختارين الفقرات ذوات معاملات الارتباط الجزئية الكبرى . هذا الاختبار لا يضع فى الاعتبار صدق الفقرة الواحدة فحسب ، بل وأيضاً ارتباطها بالمجموعة الأولى من الفقرات المصححة . وهكذا تصبح المجموعة الثانية من الفقرات هى أحسن الفقرات . ويمكننا تكرار العملية الكلية ، إذا أردنا فنصحح المجموعة الثانية من الفقرات فى الاختبارات ، ونحدد ما تزيده كل فقرة على صدق المجموعة الثانية :

والطريقة الثالثة :

لتقرير أفضل مجموعة من الفقرات يبدأ الباحث من المجموعة الكلية الموجودة من الفقرات ، مستبعداً الفقرات غير الصالحة . ذلك طبعاً بعد تحديد صدق كل فقرة . ويعد مفتاح تصحيح الاختبار الكامل . وهذا المفتاح يمكن أن يعد قبل إجراء الاختبار فى حالة اختبارات المعرفة أو القدرة . أما فى مقاييس الميول ، والتوافق ، وتاريخ الشخصية وما يشابهه ، فيعد المفتاح على أساس صدق الفقرات . وبعد إعداد المفتاح تصحح كل ورقة إجابة بواسطته . ثم نحصل على معاملات إرتباط كل فقرة بالدرجة الكلية على الاختبار . كما نحسب معامل الارتباط بين الدرجة الكلية والمحك .

وبهذه الطريقة ، يمكننا أن نحسب الارتباط الجزئى لكل فقرة فى المحك عندما تصبح الدرجة على الاختبار الكلى مقداراً ثابتاً . فإذا أردنا اختصار المقياس أمكننا حذف الفقرات ذوات معاملات الارتباط الجزئية الصغرى .

ويمكن استبعاد الفقرات باتباع أكثر من منهج . فمثلا تستبعد الفقرات التى لا يبدو أن لها قيمة ويصحح الاختبار المختصر بالنسبة لكل فقرة . ويحدد صدق درجته ، وارتباط كل فقرة بالدرجة على الاختبار المختصر . ويجعل الدرجة الكلية على الاختبارات المختصرة مقدارا ثابتا تحسب معاملات الارتباط الجزئية مع المحك وهى الأساس فى استبعاد الفقرات الأقل صلاحية .

وتواجهنا من جديد مشكلة عدد الفقرات التى نحذفها من المجموعة الأصلية من الفقرات . وفى بعض الحالات يتحدد هذا العدد بالطول المرغوب أن يكون عليه الاختبار النهائى . وتفيد معاملات الارتباط الجزئية ، فى الطريقة الثانية والثالثة ، فى تعيين الفقرات التى ستستبعد . وفى حالات أخرى لا يكون هناك مثل هذا الاعتبار العملى ، وهنا يمكن إنقاص عدد الفقرات الكلية فى خطوات متتابعة . وفى كل خطوة ، يحسب معامل

الارتباط بين الدرجة على الاختبار الكلى (بعد اختصاره) ودرجة المحك .
ونظل تستبعد فقرات ما دام معامل الارتباط بين الدرجة الكلية والمحك يتزايد . ونتوقف فى الاختصار عند ما يتناقص هذا المعامل .

استخدام الإتساق الداخلى

نركن إلى فحص الإتساق الداخلى للفقرات عند ما نكون بصدد اختبار فقراته متجانسة ، بمعنى أن كل فقرة تهدف إلى قياس نفس الوظيفة التى تقيسها الفقرات الأخرى . فنحدد ارتباط كل فقرة بالدرجة الكلية عن مجموعة الفقرات التى يتكون منها الاختبار . وهذا النوع من التحليل ، والذى يمكن القيام به فور إجراء الصورة الأولية من الاختبار على مجموعة تجريبية ، يقدم لنا معلومات ذات قيمة فى اختيار فقرات الاختبار وفى مراجعة تأليف الفقرات . وسنعالج استخدام بيانات الاتساق الداخلى فى هذين المجالين .

نفيد من بيانات الاتساق الداخلى فى اختيار الفقرات بطريقة معقدة يحسن معها استخدام صدق الفقرات فى نفس الغرض . ويرتبط الاتساق الداخلى للفقرات المفردة بتجانس الإختبار الكلى . فكلما ارتفعت درجات الاتساق الداخلى زاد تجانس الاختبار ولكن ما هى درجة التجانس التى نرتضيها للاختبار ؟ قبل الإجابة عن هذا السؤال يجب أن ننبه إلى أن التجانس ليس هدفا فى ذاته . فهو يعنى صدق درجة الاختبار ، وصدق درجات بطارية من الاختبارات ، أى مدى وضوح تفسير الدرجة على الاختبار من الناحية التحليلية .

فالفقرة ذات الارتباط المنخفض جداً بالدرجة الكلية عن الاختبار فقرة لا يمكن الاعتماد عليها وهى غالباً تقيس وظيفة تختلف تماما عن تلك التى تقيسها بقية فقرات الاختبار . وعلى هذا فالفقرات التى ترتبط ارتباطا

شديد الانخفاض أو سالباً مع الاختبار ككل ، فقرات لا يمكن الاعتماد عليها ويجب أن تستبعد تماما أو على الأقل أن تعدل وتجرب من جديد . فإذا كنا نهتم بتكوين اختبار نعطى عنه درجة مفردة لها معنى موحد ، أى درجة تتفق مع اسم وظيفة معينة ، كان علينا أن نستبعد الفقرات الأقل اتساقا ؛ ولكن هل كل الفقرات الأكثر اتساقا صالحة ؟

يجب أن يكون للفقرة ذات الاتساق مع الاختبار درجة عالية من الثبات . ويجب أن تتداخل مع الوظائف التى تقيسها الفقرات الأخرى بالاختبار ، ما دام الاختبار متماسكا فيه وحده . والاختبار الذى يقيس وظيفة واحدة ويعطى درجة واحدة يجب أن يغطى كل جوانب هذه الوظيفة .

فإذا افترضنا أن الصورة الأولية من اختبار افترض فيه قياس مدى المعرفة فى العلوم الطبيعية ، تتكون من ٢٠٠ فقرة وكانت ٨٠ فقرة تتعلق بالطبيعة و ٨٠ تتعلق بالكيمياء ، و ٢٠ فقرة تتعلق بالجيولوجيا و ٢٠ بالفلك . فإذا نظرنا إلى أحد جوانب هذه المعرفة توقعنا أن معامل الارتباط الداخلى بين الفقرات فى قسم الطبيعة مثلا يزيد عن معامل الارتباط بين فقرات قسم الطبيعة وفقرات أى قسم آخر . ولكننا نتوقع أيضاً أن قسمى الطبيعة والكيمياء لسيادتهما فى الاختبار (١٦٠ فقرة تخصهما من ٢٠٠ فقرة فى الإختبار كله) ، أن معامل الارتباط بين الفقرة فى أحدهما وبين الاختبار يرتفع . فإذا اخترنا ١٠٠ فقرة من فقرات الاختبار كله على أساس ارتفاع الاتساق الداخلى ، وجدنا أن هذه الفقرات محصورة فى قسمى الطبيعة والكيمياء . وهذا يضيق المدى الذى يفترض أن يغطيه الاختبار الذى يحقق الغرض الأصلى من إنشائه . وهكذا نرى أنه عند بحث الاتساق الداخلى يلزمنا أن نتحدد بالإطار الذى يضعه التخطيط الأصلى والمواصفات الأولى للاختبار .

وفحص الفقرات على أساس الاتساق الداخلى يتعلق بثبات وتجانس الدرجة على الاختبار الكلى الناشئ . وهكذا نصل فى نهاية الأمر إلى أن بيانات الاتساق الداخلى تفيدنا كوسيلة لاكتشاف واستبعاد الفقرات ذات الارتباط المنخفض أكثر مما تفيدنا فى اختيار الفقرات الصالحة .

والمشكلة التى تواجهنا هنا هى تقرير مدى اتساع محتوى الاختبار والوظيفة التى يقيسها حتى يصبح متجانساً . ويفترض منطقياً أن درجة الاختبار الكلى عند ما تمثل وظيفة ومحتوى محدوداً معروفاً جيداً تكون أفضل . وعلى هذا عند تكوين اختبار يقيس وظيفة مركبة من عدة اختبارات فرعية كل منها يقيس ميداناً من ميادين هذه الوظيفة ، يحسن بنا أن نربط درجة الفقرة من نوع معين بالاختبار الفرعى الذى يشملها والذى يتكون من فقرات من نفس النوع . وهنا تعطى درجة واحدة عن كل اختبار فرعى .

فبيانات الاتساق الداخلى يجب أن تستخدم أولا وقبل كل شىء فى استبعاد الفقرات غير الصالحة . ويمكن أن تفيد هذه البيانات فيما يتعلق بصعوبة الفقرة وبوضع تخطيط لمواصفات الاختبار ، وفى اختيار الفقرات الأصلح فى كل قطاع متخصص من المحتوى أو الوظيفة المقاسة . وفضلا عن هذا فإن بيانات الاتساق الداخلى فى تحليل الفقرات تعد مصدراً قيما للمعلومات يرشد مؤلف الاختبار أثناء مراجعة فقرات اختباره .

الفصل التاسع

المعايير

- مقدمة .
- الحاجة إلى المعايير .
- الصفر المطلق .
- المعايير والتقنين .
- الدرجات المحولة .
- أنواع المعايير .
- تحويل المعايير .
- النسب .
- الصفحات النفسية .
- كيف نستخدم المعايير .

الفصل التاسع

المعـــايير

مقـــدمة :

القياس والتقييم فى التربية وعلم النفس علم تعترضه عدة صعوبات أهمها أن الكثير من الصفات الإنسانية ليس خاضعاً للقياس بطريقة مباشرة ميسورة . ففى حين نجد أنه يسهل قياس الطول لأنه كمى ولأنه يزيد بوحدات منتظمة نسميها السنتيمترات وكذا الوزن والعمر الزمنى . نجد أن الكثير من الصفات الإنسانية الأكثر تعقداً مثـــل سمات الشخصية ، والخصائص العقلية ، والتحصيل ، ليست لها مقاييس متفق عليها موضوعية دقيقة دقة وموضوعية مقاييس الطول والوزن والسن .

وقد حدث منذ عصور طويلة إن كان قياس الطول والوزن والأيام والشهور يصادف مشاكل متعددة فى القياس فلم تكن هناك مستويات للمقارنة ؛ ولا وحدات متفق عليها للقياس فى هذه الميادين . وتواجهنا ، فى أيامنا هذه نفس المشاكل ، من إقرار وحدات لقياس السلوك البشرى فى المجالات التى لم تحدد لها معايير مماثلة . وليست المشكلة هى إيجاد الوحدات بقدر ما هى البحث عن وحدات ذات دلالة ومعنى يمكن تفسيرها ومقارنتها . والمشكلة أيضاً هى وضع مستويات بها نقارن ونقيس ونقيم ونفسر ونعلل السلوك الإنسانى . هذه الوحدات هى المعايير .

الحاجة إلى المعايير

إذا رجعنا إلى الفصل الثانى (مثال درجات التلاميذ فى تفسير الدرجة الخام) وجدنا أن الدرجة الخام ومثيلاتها النسبة المئوية لعدد الناجحين إلى

العدد الكلى ، أو النسبة المئوية لعدد الفقرات الصحيحة إلى عدد فقرات الاختبار ، ليس لها فى ذاتها معنى أو دلالة . فهى لا تحدد مركز الشخص بالنسبة لمجموعته أو بالنسبة لعينة التقنين ، ولهذا لا بد من الرجوع إلى معيار يحدد معنى هذه الدرجة أو النسبة المئوية . فيدلنا على مركز الشخص بالنسبة للمجموعة وهل هو متوسط أو فوق المتوسط أو أقل من المتوسط . وما مدى بعده عن متوسط المجموعة التى ينتمى إليها ؟ وما هو وضعه بالنسبة لمن يكافئونه من أفراد عينة التقنين ؟

ذلك أن الدرجة الخام أو النسبة المئوية للإجابات أو للأفراد تعجز عن إعطاء أى تفسير . فإذا كانت النسبة هى ٥٠ ٪ إجابات صحيحة من فقرات اختبار معين فإنها قد تساوى ٣٠ ٪ إجابات صحيحة من فقرات اختبار آخر أصعب منه ، وهى قد تساوى فى نفس الوقت ٨٠ ٪ فى اختبار آخر أسهل منه . وقد حدثت تجارب من هذا النوع أعطيت فيها للأفراد اختبارات تقيس وظيفة معينة ولكن مستوياتها مختلفة ، فكانت درجة الفرد تنخفض كلما ارتفع مستوى صعوبة الاختبار . وعلى هذا لا يمكننا أن نقول إن فلانا أعلى من فلان فى قدرة ما لأن الأول حصل على درجة فى أحد الاختبارات تعلو الدرجة التى حصل عليها الآخر فى اختبار آخر حتى لو كانا يقيسان نفس الوظيفة .

وهكذا لا يكون للدرجة الخام ولا للنسبة المئوية دلالة فى حد ذاتها بل تحتاج إلى معيار يكسبها معنى . يجب إذن أن نبحث عن نوع أفضل من الوحدات لنعبر بها عن نتائج الاختبار بطريقة أفضل من العد الخام للفقرات الصحيحة أو لوحدات الدرجات والنسبة المئوية . ونحن نفضل أن تتمتع هذه الوحدات ، أى المعايير ، بالخصائص الآتية :

١ ــ أن يكون للدرجة الواحدة معنى موحد من اختبار إلى آخر . حتى

نوفر أساساً يمكن به أن نقارن درجات الأفراد فى اختبارات مختلفة :

٢ ـ أن تكون هذه الوحدات متساوية بحيث أن العدد المعين من الوحدات ، وليكن ١٠ درجات ، على جزء من اختبار يدل على نفس الشيء الذى تعنيه ١٠ درجات على جزء آخر من نفس الاختبار : بحيث يمكننا أن نقول إن الفرد الأول يزيد عن الفرد الثانى بـ ١٠ درجات مثلا ، وهذه الزيادة تساوى زيادة الشخص الرابع عن الشخص الثالث .

٣ ـ أن تكون هناك نقطة صفر حقيقية ، صفر مطلقة ، تعبر عن لا شيء من الصفة التى نقيسها ، بحيث نستطيع أن نقول إن درجة معينة ضعف درجة أخرى وأن ثالثة تعتبر ثلثى درجة رابعة .

الصفر المطلق :

والأنواع المختلفة من المعايير التى طورتها الاختبارات تمثل تقدماً ملحوظاً نحو الهدفين الأولين . ولكن النقطة الثالثة يحتمل أن لا نصل إليها أبداً فى السمات التى يخصها المقياس النفسى والتربوى . فلسنا هنا بصدد مقاييس من نوع مقاييس الأوزان بحيث نضع سنجة ٥ أرطال فى كفة من كفتى الميزان ونضع أمامها ٥ أرطال من مادة معينة تزنها تماماً . وهنا حينما نقول « لا يزن شيئاً » نعنى فعلا « أنه لا يزن شيئاً » ويمكننا أن نضيف وحدات الوزن معاً ، وليس لدينا هذا النوع من نقطة الصفر أو هذه الطريقة من الجمع معاً فى حالة القياس النفسى والتعليمى : فإذا جمعنا شخصين ضعيفى العقل معاً فلن يعادلا عبقرياً كما أن زوجاً من الضعاف فى الهجاية لا يفوق مجيد الهجاية .

المعايير والتقنين

قلنا إن عمليات التقنين هى الخطوات التجريبية التى يمر بها المقياس فى صورته النهائية عند إجرائه على عينات من الجماعات التى سيطبق عليها فيما بعد ذلك بغرض اشتقاق المعايير . وعلى هذا فنحن نشتق المعايير من عينة التقنين

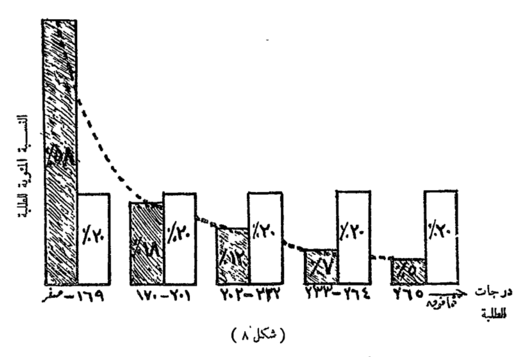
Standardization Sample التى تمثل المجتمع الأصلى المدروس Population، فتكون الدرجات المحولة للدرجات الخام لأفراد هذه العينة مصدراً للمعايير . أى أننا عنـــد إجراء الاختبار نقارن المفحوص بمن يكافئه من أفراد عينة التقنين .

فإذا تساوت درجة المفحوص مع درجة فرد فى عينة التقنين ، أو مع متوسط إحدى فئات هذه العينة ، كان المركز النسبى للمفحوص فى الجماعة التى ينتمى إليها معادلا لمركز فرد معين فى هذه العينة ، أو لمركز فئة معينة فى هذه العينة . وهكذا فإننا نقارن درجة الفرد الخام بتوزيع الدرجات الخام لأفراد عينة التقنين ، فتتحول درجته إلى معيار من نفس نوع معيار الاختبار . وبذا تسهل مقارنة الفرد بالمستوى المتوقع له بما أنه عضو فى جماعة معينة ، أى أننا نقارنه بمتوسط هـــذه الجماعة أو بمتوسطها وانحرافها المعيارى كما سيتضح لنا فيما بعد .

وعلى ذلك فالمعايير ليست إلا نتائج إجراء الاختبار على عينات التقنين . فإذا كانت هذه العينات ممثلة للمجتمع الذى نود قياسه بالاختبار ، صلحت هذه المعايير ما دام المجتمع المراد دراسته قد مُثل تمثيلا صادقاً عند اختيار عينة منه . وإلا أصبحت المعايير غير ذات قيمة أو ذات قيمة محدودة . ومن هنا وجب على كل مستخدم للاختبار أن يجريه فقط على أفراد أو مجموعات مثلت فى عينات التقنين . وعند نقل الاختبار من بيئة إلى أخرى يجب تقنينه على عينات تعد موازية لعينات التقنين الأصلية لاشتقاق المعايير الجديدة ؛ حتى يمكن قياس جماعات مقصـــودة تعد مكافئة للمجتمع الأصلى الذى كان يراد قياسه بالاختبار الأصلى قبل نقله . انظر الشكل (٨) .

وما دامت عينة التقنين عماد المعايير ، لذا يحسن أن نوضح الخصائص التي يجب أن تتوفر فيها حتى يصح قياس وتقييم من نود قياسه وتقييمه : ويمكن تلخيص أهم الخصائص التي يجب أن تتوفر في عينة التقنين الجيدة فيما يلي :

١ ـ يجب أن تمثل عينة التقنين المجتمع الأصلي المراد دراسته تمثيلا صادقاً من حيث الحجم والتركيب والنسب ومستوى القـــدرة أو الصفة المطلوب قياسها . ولا يتوفر هذا إلا بقيام الباحث ، قبل اختيار عينة التقنين ، بدراسة هذا المجتمع الأصلي دراسة مسحية Survey study كي يحدد ما سبق ذكره وكي يعرف بالتحديد حجم العينة التي يجب أن يختارها وخصائصها ، وحجم العينة التي سيختارها بالفعل وخصائصها . ذلك أنه دائماً ما تحول صعوبات عملية دون اختيار أمثل عينة . ونتيجة لهذا يتعين على واضع الاختبار أن يذكر صراحة نقائص عينته وحدود معاييره .



(شكل ٨)

يوضح الشكل أهمية اشتقاق معايير محلية . والمستطيلات المظللة تمثل توزيع درجات ٤٠٢٩٩ طالب في ٢٠٣ جامعة بحسب المعايير القومية . ويلاحظ أنها تختلف دائما عن المستطيلات غير المظللة التي تمثل توزيع ٦٤٦ طالبا بجامعة شيكاغو بحسب المعايير المحلية .

٢ — كلما كبر حجم العينة ، أى كلما زاد عدد أفرادها ، زاد الاعتماد على نتائج إجراء الاختبار عليها . وقد سبق أن بينا أن صدق الاختبار وثباته يتوقفان على حجم العينة وتمثيلها ؛ ولكن يجب أن ننبه إلى أنه ليست هناك أحجاماً معينة يجب أن تكون عليها عينة التقنين الجيدة ، فإن هذا يتوقف على عدة عوامل أهمها :

(أ) حجم المجتمع الأصلى المقصود دراسته .

(ب) طبيعة الاختبار ونوع الوظيفة المقاسة والغرض المنشود من عملية القياس ، فاختبار الفصل يقنن على الفصل واختبار المدرسة يقنن على المدرسة . بينما يقنن اختبار الذكاء على أكبر عينة ممكنة ويقنن الاختبار الإكلينيكى عادة على عينة أصغر ؛

(ح) إمكانيات الباحث ومدى تعاون أفراد العينة أو المسئولين عنها مع الأخصائى الذى يجريه أو الهيئة المشرفة عليه .

٣ — يجب أن يتساوى متوسط وتشتت أفراد العينـــة مع متوسط وتشتت أفراد المجتمع الأصلى .

٤ — كلما صغر حجم القطاع المراد قياسه تمكن الباحث من تمثيله فى العينة وصدقت معاييره .

ويمكن الرجوع إلى الفصل الثانى لمعرفة طرق اختيار العينات .

الدرجات المحولة

ولما كانت الدرجة الخام ومثلها النسبة المئوية ، لا تعنى فى حد ذاتها شيئاً كان لابد من تحويلها حتى تصبح ذات دلالة للحاصل عليها .

وتلك الدرجات المحولة Trensformed score هى التى نقصدها عند الكلام عن المعايير . وتخدم المعايير غرضين :

١ — فهى تحدد مركز الفرد بالنسبة لعينة التقنين .

٢ — تمكننا من مقارنة مركز الفرد على مقياس بمركزه على غيره — فإذا

حصل فرد فى مقياس مفردات على ٤٠ درجة وفى مقياس قدرة ميكانيكية على ٢٢ درجة ، ففى أيهما يكون أقوى ؟ ولا يتيسر ذلك إلا بالرجوع إلى معايير كل مقياس لمعرفة مركز الفرد فى كليهما ثم مقارنة مركزه على هذا بمركزه على ذاك .

ولا يمكن أن نعطى معنى لدرجة خام إلا بالرجوع إلى نوع معين من الجماعة أو الجماعات . فالدرجة لا تكون عاليــة أو منخفضة ولا جيدة أو رديئة ولكنها تكون أعلى من . . . أو أقل من . . . ، أو أحسن من . . . أو أسوأ من . . . وهناك طريقتان عامتان نربط بهما درجة الفرد بإطار عام . والطريقة الأولى هى أن نقارنه بسلسلة متدرجة من الجماعات وأن نبحث عن الجماعات التى يتكافأ الفرد مع أفرادها بما أنه قد حصل على نفس متوسط درجاتهم . وكل جماعة فى هذه السلسلة عادة تمثل فرقة دراسية معينة أو عمراً زمنياً معيناً . والطريقة الأخرى هى أن نحدد موضعه فى جماعة معينــة بالإشارة إلى النسبة المئوية من الجماعة التى يمتاز بها أو بالرجوع إلى متوسط الجماعة وانحرافها المعيارى :

أنواع المعايير

وعلى هذا نجد ٤ أنماط أساسية فى تفسير درجة الفرد :

نوع المعيار	نوع المقارنة	نوع الجماعة
١ – معايير السن	مقارنة الفرد بالجماعة التى يكافئها	جماعات متتابعة فى العمر
٢ – « الفرقة الدراسية	مقارنة الفرد بالجماعة التى يكافئها	جماعات متتابعة فى الفرقة الدراسية
٣ – معايير المئين	النسبة المئوية من الجماعة التى يمتاز عنها الفرد	جماعات من السن أو الفرقة الدراسية التى ينتمى إليها الفرد
٤ – معايير الدرجة المعيارية	عدد الانحرافات المعيارية التى يزيد أو يقل بها الفرد عن متوسط الجماعة	جماعات من نفس السن أو الفرقة الدراسية التى ينتمى إليها الفرد

جدول (٥) المقارنة بين المعايير الشائعة .

معايير العمر :

أى سمة تنمو مع زيادة السن يمكننا أن نعد لها معايير عمر . ومعيار العمر هنا هو القيمة المتوسطة لهذه الصفة لأشخاص من نفس العمر . ولكى نوضح هذا نضرب المثال التالى :

لو اخترنا عينة ممثلة لأطفال من عمر ٥ سنوات . ووزنا كل واحد منهم ثم حسبنا متوسط وزن كل أفراد هذه العينة . فإن هذا المتوسط يكون معيار الوزن لهذا السن . ونلاحظ هنا أن هذا المعيار ليس إلا مجرد قيمة متوسطة ، وليس نموذجاً . كما وأنه لا ينتظر أن يكون وزن كل فرد من هذا السن هو هذا الوزن المتوسط .

وبنفس الطريقة يمكننا تقدير الوزن المتوسط للأطفال من عمر ٣ سنوات ، ٤ سنوات ، ٦ سنوات . وبعد هذا يمكننا تمثيل كل متوسط بنقطة على رسم بيانى ونوصل النقط فيصبح لدينا منحنى .

والنقط على هذا المنحنى محسوبة من أوزان السنوات الكاملة أى أن العينة فى كل سن كانت كلها تصل إلى حد هذا السن مثلا كلهم سن ٣ سنوات تماماً ، ٤ ، ٥ ، ٦ ، . . . الخ . ولكن يمكننا اعتبار هذا المنحنى مستمراً Continuous أى يمكننا تقدير أوزان الأطفال الذين تقع أعمارهم بين حدى عمرين . مثلا طفل وزنه ١١ كجم ووجدنا أن متوسط وزن طفل الثالثة هو ١٠ كجم ومتوسط وزن طفل رابع هو ١٢ كجم لتبين لنا أن طفلنا يعادل فى وزنه الطفل المتوسط من عينة التقنين الذى سنه ٣ سنوات و ٦ شهور .

ومن الناحية الأخرى ، إذا علمنا عمر الطفل بالتحديد وليكن ٥ سنوات ووزناه فوجدنا أنه يزن ١٢ كجم لعلمنا أن وزنه أقل من متوسط وزن عمره . وإذا كان طفل ثالث سنه ٣ سنوات ووزنه ١٢ كجم كان ، بالتالى ، بديناً .

هذه الطريقة فى تقدير أوزان الأفراد أو أى سمة تنمو بتقدم العمر ، هى طريقة سهلة وعملية . ولكن لها عدة عيوب :

١ ــ وحدات معايير الأعمار غير متساوية . فإن الزيادة فى الوزن بين ٣ سنوات و ٤ سنوات لا تعادل الزيادة فى الوزن بين سن ١١ ، ١٢ سنة . وبالمثل بالنسبة لكل عمر على مقياسنا هذا . فإذا نظرنا إلى سن ١٥ سنة ، مثلا ، وجدنا أن الزيادة فى الوزن لا تتناسب مع الزيادة فى العمر تماماً . وبعض الصفات الأخرى غير الوزن كثيراً ما نجد أن زيادة الصفة تتوقف عند سن معينة . فلا يزيد الطول بعد سن العشرين مثلا وكلما اقتربنا من هذا السن قلت سرعة النمو . إن هذه الوحدات ، إذن ، ليست ذات معنى موحد .

والجدول (٦) جدول لتوزيع الأفراد حسب أعمارهم (العمود الرأسى) الذين حصلوا على درجات معينة فى اختبار للمفردات (الصف الأفقى) . وعلى هذا فكل خانة ثقيلة الخطوط تقابل درجة هى معيار الأفراد من سن هو رقم العمود الذى يشمل هذه الخانة . ويكون معيار سن ٧ هو الدرجة ٢٨ وهكذا .

هذا النوع من المعايير لا يفيد فى تقدير سمات الأفراد المتطرفين فى التفوق أو التخلف .

كذلك فإن منحنيات النمو للسمات المختلفة لا يمكن مقارنتها مع بعضها فإن سرعة النمو تختلف من صفة لأخرى لدى الفرد كما تختلف بين الأفراد فى الصفة الواحدة .

فهناك وظائف تستمر فى النمو حتى سن متقدمة وأخرى يتوقف نموها فى سن مبكر نسبياً .

(جدول ٦)

معايير العمر فى اختبار للمفردات

العمر بالسنة	٦	٧	٨	٩	١٠	١١	١٢
٢٠	X						
٢١	XXX						
٢٢	XXXX	X					
٢٣	XXXXX	XX	X				
٢٤	XXXXXX	XX	X				
٢٥	XXXXXXX	XXX	XX				
٢٦	XXXXXX	XXXX	XXX	X			
٢٧	XXXXX	XXXXX	XXXX	XX	X	X	
٢٨	XXXX	XXXXXX	XXXXX	XX	XX	XX	
٢٩	XXX	XXXXX	XXXXXX	XXX	XX	XX	
٣٠	X	XXXX	XXXXXXX	XXX	XXX	XXX	X
٣١		XXX	XXXXXX	XXXX	XXX	XXX	XX
٣٢		XX	XXXXX	XXXXX	XXXX	XXXX	XXX
٣٣		XX	XXXX	XXXXXX	XXXXX	XXXXX	XXX
٣٤		X	XXX	XXXXX	XXXXXX	XXXXX	XXXX
٣٥			XX	XXXX	XXXXXXX	XXXXXX	XXXX
٣٦			X	XXX	XXXXXX	XXXXXXX	XXXXX
٣٧			X	XXX	XXXXX	XXXXXXX	XXXXXX
٣٨				XX	XXXX	XXXXXXX	XXXXXX
٣٩				XX	XXX	XXXXXX	XXXXXX
٤٠				X	XXX	XXXXX	XXXXXXX
٤١					XX	XXXX	XXXXXX
٤٢					XX	XXXX	XXXXXX
٤٣					X	XXX	XXXXXX
٤٤					X	XXX	XXXXX
٤٥						XX	[illegible]
٤٦						XX	XXXX
٤٧						X	XXX
٤٨							XXX
٤٩							XX
٥٠							X

علامة (X) تمثل تكرار الدرجة ، والخانة المحددة بخط ثقيل (سميك) تقابل معيار العمر أو متوسط درجة العمر الذى تتوسطه . فمثلا المعيار أو الدرجة المتوسطة لعمر ٦ سنوات هى ٢٥ ولسن ٧ سنوات هى ٢٨ ولسن ٨ سنوات هى ٣٠ . . . . الخ

معدل سرعة النمو في سن معين يختلف عنه في سن آخر فالطفل من عمر ٣ سنوات عندما يبلغ وزنه ١٢ كجم مثلا تكون درجة بدانته واضحة أكبر من بدانة طفل ١١ سنة الذى يصل وزنه إلى متوسط وزن أطفال الثانية عشرة . هذه الملاحظة هى التى أدت إلى استخراج I.Q وغيرها من النسب التى سنتعرض لها فيما بعد . هذه النسب تتلافى الفروق فى معدل سرعة النمو . وعلى أى حال فإن مشكلة عدم تساوى وحدات العمر ستظل قائمة . والسمات التى لا يكون نموها مضطرداً مع تقدم العمر مثل حدة البصر لا يمكن تمثيلها بوحدات العمر . وكذلك المهارات التعليمية مثل التى تنمو مع تقدم الفرد فى دراساته فيكون من الأنسب أن نقارنها بمعايير الفرقة الدراسية .

وقد استخدمت معايير العمر على نطاق واسع فى كثير من اختبارات الذكاء العام والقدرات العقلية . وهنا كان يحدد عمر عقلى **Mental Age** وعمر زمنى **Chronological Age** .

العمر العقلى

قلنا إن أى صفة تنمو بتقدم العمر يمكن اشتقاق معايير عمر لها . والنمو العقلى مضطرد ، إلى حد ما ، مع العمر الزمنى . وقد كان هذا المفهوم دافعاً فى مراجعة مقياس بينيه – سيمون سنة ١٩٠٨ ، إلى وضع فقرات يتناسب مستوى صعوبتها مع مستوى قدرة كل سن . وعلى هذا يطبق الاختبار ، الذى يستخدم معايير العمر العقلى ، على عينات من مختلف الأعمار تمثل المجتمع الأصلى بالأعمار المقصودة . ويحسب متوسط درجات أو مستوى إجابات كل سن على حدة ، ويعتبر هذا المتوسط هو معيار العمر العقلى لهذا السن .

فإذا حصل فرد على درجة معينة ، فإننا نبحث عن العمر الزمنى الذى كان متوسط درجات أفراد العينة فيه مساوياً لهذه الدرجة . فإذا كانت

درجة الفرد الخام هى ٥٠ مثلا فإننا نبحث فى معايير الأعمار عن الفئة العمرية التى كان متوسط درجات أفرادها هو ٥٠ ، وليكن عمرها الزمنى هو ١٥ ، فيصبح العمر العقلى لهذا الفرد هو ١٥ . أى أن العمر العقلى للفرد هو العمر الذى تتساوى فيه درجة هذا الفرد مع متوسط درجات أفراد هذا العمر من عينة تقنين اختبار الذكاء المعين .

هذا المعيار يقوم ، كما سبق أن اتضح لنا ، على أساس أنه كلما زاد عمر الفرد نما عقلياً . أى أن الأفراد من سن معين يستطيعون الإجابة عن فقرات لا يستطيعها أفراد من سن أقل . ولكن هذا الافتراض غير صحيح ، إذ أن الفرد الواحد قد يستطيع إجابة أسئلة صعبة وضعت فى سن ١٤ مثلا ولا يستطيع فى نفس الوقت أن يجيب عن فقرات أسهل وضعت فى سن ١٣ . هذا التشتت فى إجابات المفحوصين هو الذى دعى إلى فكرة العمر القاعدى ، وهو أعلى عمر يستطيع فرد معين أن يجيب إجابة صحيحة عن كل فقراته . وهو الذى دعى أيضاً إلى فكرة العمر الأقصى ، وهو العمر الذى لا يستطيع فيه المفحوص أن يجيب إجابة صحيحة عن أى فقرة .

وبين العمر القاعدى والعمر الأقصى يجيب فيها المفحوص عن فقرات فيحصل على درجات تحول إلى شهور . وتجمع هذه الشهور إلى العمر القاعدى وتكون النتيجة هى العمر العقلى للفرد . وبقسمة هذا العمر العقلى على العمر الزمنى لنفس المفحوص يخرج لنا كسر إذا ضرب فى مائة حصلنا على نسبة ذكاء الفرد .

$$\text{أى أن نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

$$= \frac{\text{ع . ع M. A.}}{\text{ع . ز C. A.}} \times ١٠٠$$

ويمكن حساب العمر العقلى ونسبة الذكاء باختبارات غير مقسمة إلى مستويات عُمرية . ذلك بأن يحصل الفرد على درجة فى الاختبار كله نقارنها بمتوسط درجة من يساوونه فى العمر من أفراد عينة التقنين ، فإذا تساوت الدرجتان كان الفرد متوسطاً . وإذا زادت درجة الفرد عن متوسط درجات عمره كان متفوقاً ، وإذا حدث العكس كان متخلفاً (انظر فصل قياس الذكاء لمعرفة مزايا وعيوب العمر العقــلى والطرق المختلفة فى حساب نسبة الذكاء) .

عدم ثبوت وحدات العمر العقلى :

لا تظل وحدات العمر العقلى ثابتة بل تتناقص بتقدم العمر فالسنة تأخر فى سن ٤ تعادل حوالى ٣ سنوات تأخر فى سن ١٢ وهكذا فطفل عمره الزمنى ٤ والعقلى ٣ يكون متخلفاً عاماً واحداً عن متوسط سنه وتكون نسبة ذكائه $\frac{٣}{٤}$ × ١٠٠ = ٧٥ . ونسبة الذكاء ثابتة تقريباً فعندما يكون عمره ١٢ يكون عمره العقلى = $\frac{\text{ع . ع .}}{١٢}$ = ٧٥ ∴ ع . ع = ٩ سنوات أى أصبح متأخراً ثلاث سنوات عن متوسط سنه . وسبب هذا أن سرعة النمو العقلى تختلف ففى الطفولة تزيد كثيراً جداً عنها فى الأعمار الأكبر .

ويقف نمو الذكاء عادة عند سن ١٥ — وإذا طبقنا تلك المعادلة فستقل نسبة ذكاء الشخص كلما كبر عن سن ١٥ . لهذا يكتفى بحساب عمر زمنى ١٥ سنة[٦] لأى شخص عمره أكثر من ١٥ ســنة — ونعوض فى المعادلة كالمعتاد . وميزة العمر العقلى هى أنه واضح ومفهوم حتى لدى العامة .

عيوب العمر العقلى :

١ — عدم ثبات تباين الأفراد فى الأعمار المختلفة . فيزداد التباين كلما تقدم العمر ذلك لتناقص وحدات العمر إذ أن سنة فى عمر زمنى ٤ تعادل

حوالى ٣ سنه فى عمر زمنى ١٢ وعلى هذا يزيد التباين فى الأعمار الكبيرة ويقل فى الأعمار الصغيرة . وتتلافى نسبة الذكاء هذا العيب .

٢ ــ يفقد العمر العقلى وضوحه ومدلوله فى حالة الراشـــد المتفوق والراشد المتقدم فى العمر . فالراشد المتفوق ذو العمر الزمنى ١٥ والعمر العقلى ٢٠ لا يعنى أنه يساوى ذكاء عمر عقلى لشخص متوسط فى عمر زمنى ٢٠ إذ أن هذا الأخير لا يزيد عمره العقلى عن ١٥ حيث لا ينمو الذكاء بعد سن ١٥ وهكذا يفقد العمر العقلى مدلوله فى هذه الحالات .

٣ ــ لا نستخدم درجات العمر إلا فى السمات والقدرات التى تنمو بتقدم السن حيث تزداد مثل الذكاء أما تلك السمات التى لا تتأثر بتقدم العمر فيستحيل استخدام درجات العمر فيها ومن أمثال ذلك معظم سمات الشخصية كالانبساط والانطواء . الخ .

أما فى ميدان التربية نتيجة لطبيعة الاختبارات وللنقد الموجه إلى مفهوم العمر العقلى ، فيغلب استخدام معايير الفرقة الدراسية **Grade Norms** وهى قريبة فى أساسها من معايير العمر .

معايير الفرقة الدراسية

لا تختلف معايير الفرقة الدراسة **Grade Norms** كثيراً عن معايير العمر ، فمعيار الفرقة هو متوسط درجات أفراد الفرقة فى الاختبار المعين . وهنا نقارن الفرد المعين بمتوسط درجات الفرقة التى ينتمى إليها . أو نقارن الفرد بفرقة أخرى اقتربت درجته من متوسطها فيعد فى هذه الفرقة بالنسبة للمادة المقاسة . أو نقارن الفرد بمن فى سنه من أفراد فرقته . وهنا نستبعد المتطرفين من الفرقة ســواء كانوا صغاراً فى السن أو كباراً عن العمر المنوالى **modal age** لهذه الفرقة .

ولمعايير الفرق الدراسية نفس عيوب معايير العمر والعمر العقلى . (ولمعرفة المزيد يمكن الرجوع إلى فصل خصائص اختبارات التحصيل) . وعموماً نجد أن هذه المعايير أسهل فى حسابها ، أسهل لأنها مبنية على مجموعات مكونة بالفعل فى النظام المدرسى أى أن الباحث لا يقوم بتقسيمها إلى مستويات كثيرة . ومفهوم « مستوى الفرقة الدراسية » واضح عن مفهوم « مستوى العمر » . ولكن المفهوم الأول خارج الفرق الدراسية يكون محدود القيمة .

المعايير المئينية

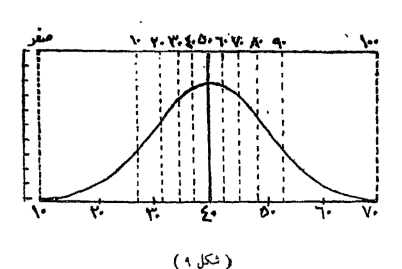
لاحظنا فى كلامنا عن معايير العمر ومعايير الفرق الدراسية أننا نقارن الفرد بمتوسط عمر معين أو فرقة دراسية معينة . ولكن من الأجدى أن نقارن هذا الفرد بالآخرين من عمره أو الآخرين من فرقته ، حتى نحدد مركزه بالنسبة لهذه الجماعة . كأن نعرف مثلا أنه يفوق ٧٥ ٪ مثلا من هذه الجماعة فى مادة معينة أو فى اختبار يشمل عددا من المواد .

وتتميز المعايير المئينية بأنها أكثر مرونة وأوسع مدى فى تطبيقها عملياً . إذا يمكن تطبيقها على أى جماعة لتحديد موضع كل فرد بالنسبة للآخرين . وهى تطبق على أوسع نطاق فى مجالات الصناعة . وعند تطبيق المعايير المئينية لاختبار منشور يجب على الفاحص أن يكملها باستخراج معايير خاصة ، بالمجموعة التى يطبقه عليها .

كأن تستخرج معايير محلية خاصة بدرجة معينة من واقع درجات تلاميذها مما يوفر فرصة تفسير درجة التلميذ الفرد على ضوء درجات مجموعته المحلية . وهذا التفسير بالطبع أجدى من مقارنة التلميذ بالمعايير المئينية المشتقة من تلاميذ الدولة كلها .

والمشكلة الأولى هى مشكلة اختيار عينة التقنين التى يجب أن تمثل فرقاً وأعماراً

مختلفة من المجموع السكانى الذى نقيسه . وهنا نحتاج إلى عمل معايير مئينية لكل فرقة دراسية على حدة أو لكل سن على حدة .



(شكل ٩)

يوضح الشكل تركز توزيع الدرجات فى الوسط وقلة التوزيع تدريجيا فى الطرفين . لاحظ أن الخط القاعدى يمثل درجات خام وحداتها متساوية وأن الخط الموازى له يمثل مئينات وحداتها غير متساوية

والمشكلة الثانية بالنسبة للمعايير المئينية هى عدم تساوى الوحدات (التوزيع فى الأطراف والتوزيع فى المنطقة المتوسطة انظر الشكل (١٠) . فالفرق بين رتبتين مئينيتين فى وسط المنحنى أصغر كثيراً من الفرق بين رتبتين مئينيتين فى طرفى المنحنى . وسيتضح هــذا بعد أن نعرف الدرجة المئينية .

الدرجة المئينية هى الدرجة التى تقع تحتها نفس النسبة المئوية من الأفراد وهى النسبة المئوية للأشخاص الواقعين تحت درجة معينة . فمثلا إذا حل ٢٨٪ من أفراد العينة أقل من ١٥ وحدة عن المقياس فإن الشخص الذى يحل ١٥ وحدة يقع فى المئين الـ ٢٨ ويرمز لها بالإنجليزية p. 28 أى أن

المئين يحدد مباشرة ترتيب الفرد فى المجموعة حيث المئين الـ ٥٠ هو المتوسط والمئين صفر يشير إلى درجة أقل من أى درجة فى المجموعة والمئين ١٠٠ يحدد درجة أعلى من أى درجة فى المجموعة . ويجب أن تفرق بين المئين وهو معيار وبين النسبة المئوية وهى درجة خام فى حاجة إلى أن تقارن بمعيار حتى يتضح مدلولها .

وفى حالة اختبارات ذكاء أقل من ١٥ سنة يصبح لكل مجموعة سن معينة مئينات خاصة بها . فهناك مثلا مئينات لسن ٧ وأخرى لسن ٨ . . . الخ . وكذلك فى حالة ما إذا كان هناك خلاف كبير بين استجابات المجموعات الفرعية Subgroups داخل المجموعة الكلية مثلا اختلاف قدرة الإناث فى القدرة الميكانيكية — وضعفهم فيها عن الذكور . ففى أمثال هذه الحالات يحسن أن توجد معايير خاصة لكل من الإناث والذكور .

ميزة المئينات :

١ — تعطى صورة صادقة لترتيب الفرد بالنسبة للمجموعة .

٢ — سهلة فى حسابها واضحه فى مدلولها .

٣ — أوسع فى انتشارها من درجات العمر إذ يمكن استخدامها بالنسبة للأطفال كما للراشدين إلى جانب أنها صالحة لأى نوع من المقاييس الشخصية أو الذكاء أو القدرات أو التحصيل .

عيوب المئينات :

١ — عدم تساوى الوحدات المئينية على منحنى التوزيع إذ تقل المسافات بين المئينات فى الوسط وتزيد كلما اتجهنا نحو الوسط . وبهذا فإن المئينات تبالغ فى تباين الأفراد فى الوسط وتقلل التباين فى الأطراف . فمثلا نجد حوالى ١٠٪ من الحالات تقع على مسافة من الدرجات الخام تقارب سبعة

أمثال المسافة التى تغطيها ١٠٪ من الحالات فى الوسط حيث نجد المسافة بين المئين صفر ، ١٠ تساوى سبعة أمثال المسافة بين المئين ٤٠ و ٥٠ وهذا يعنى أنه ليس هناك اتفاق بين توزيع الدرجات الخام وتوزيع المئينات .

٢ ـ لا يعطينا المئين مدى اختلاف الدرجة الخام عن غيرها وكل ما يعطيه هو ترتيبها فقط .

الدرجات المعيارية :

الدرجة المعيارية هى المسافة التى تبعدها الدرجة الخام عن المتوسط الحسابى معبرا عنها فى وحدات من الانحراف المعيارى Standard Deviation of the distrubution والانحراف المعيارى هذا هو مقياس من مقاييس التشتت أو التباين ويعبر عن مدى الاختلافات الفردية فى درجات المقياس

استخراج الدرجة المعيارية :

وتستخرج الدرجة المعيارية بالتعويض عن المعادلة الآتية :

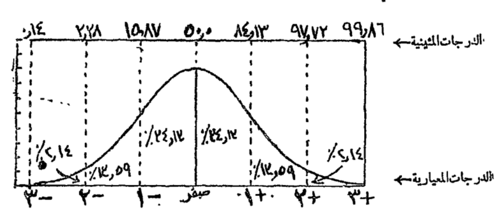
$$\text{الدرجة المعيارية} = \frac{\text{الدرجة الخام} - \text{المتوسط الحسابى للدرجات}}{\text{الانحراف المعيارى للدرجات}} =$$

وواضح أن الدرجة الخام الأقل من المتوسط ستكون إشارتها عند تحويلها إلى درجة معيارية سالبة . أما الدرجة الخام الأعلى من المتوسط ستكون موجبة الإشارة والدرجة الخام التى تعادل المتوسط ستكون مساوية لصفر درجة معيارية .

ويجب أن تحسب الدرجة المعيارية إلى كسر عشرى من رقم واحد على الأقل لكى تتمكن من التمييز بين الأفراد حيث غالباً ما تقع درجات التوزيع بين +٣ درجة معيارية ، ـ ٣ درجة معيارية حسب التوزيع لاعتدالى الذى يغلب فى الاختبارات النفسية . (انظر شكل ١٠)

وهكذا يتضح لنا أن الدرجات المعيارية تحسب على أساس قيمتى المتوسط والانحراف المعيارى بينما وجدنا من قبل أن معايير العمر والفرقة الدراسية والمئين لا تلتزم بانحراف الفرد عن متوسط جماعته إنحرافا يعبر عنه عدديا بالإنحراف المعيارى . أما الدرجة المعيارية فهى تحسب نسبة انحراف الفرد عن المتوسط (س — م) إلى انحراف أفراد الجماعة كلهم عن المتوسط (ع) . وهذه النسبة تكون موجبة الإشارة فى حالة من يعلو المتوسط وسالبة الإشارة فى حالة من يقل عنه . وطبيعى أن يتساوى عدد من يزيدون عن المتوسط مع عدد من يقلون عنه ، وطبيعى أيضاً أن تتساوى الانحرافات السالبة مع الانحرافات الموجبة . وعلى هذا يكون مجموع الدرجات المعيارية لأى توزيع هو صفرا (+٣ —٣ = صفر) . وأهم خاصية فى الدرجات المعيارية هى أن وحدة حسابها هى الانحراف المعيارى . وإذا حسبنا الانحراف المعيارى للدرجات المعيارية لأفراد أى مجموعة وجدناه مساويا واحدا صحيحا .

ومن شأن كل هذه الخصائص أن تسهل الدرجات المعيارية عمليات مقارنة المجموعات المختلفة حتى لو اختلفت متوسطاتها وانحرافات المعيارية ما دامت الواحدات ، فى الدرجات معيارية وهى الانحرافاتها المعيارية ، متساوية . وبهذا يمكن مقارنة درجات الفرد الواحد فى اختبارين مختلفين أو أفراد مختلفين فى اختبار واحد .



(شكل ١٠) لاحظ أن المئينات المقابلة للدرجة المعيارية من صفر إلى + أضعاف المئينات المقابلة للدرجة المعيارية من + ٢ إلى + ٣ مثلا

وفيما يلى توزيع مئوى للحالات فى المنحنى الاعتدالى Normal Curve بين الدرجات المعيارية المختلفة محسوب من الشكل (١٠) .

بين − ٣ ، − ٢ يقع ٢ر١٤ ٪ من الحالات
بين − ٢ ، − ١ يقع ١٣ر٥٩ ٪ من الحالات
بين − ١ ، صفر يقع ٣٤ر١٣ ٪ من الحالات
بين صفر ، + ١ يقع ٣٤ر١٣ ٪ من الحالات
بين + ١ ، + ٢ يقع ١٣ر٥٩ ٪ من الحالات
بين + ٢ ، + ٣ يقع ٢ر١٤ ٪ من الحالات .

وبهذا يكون تقابل المئينات والدرجات المعيارية على الوجه الآتى من التوزيع الاعتدالى :

− ٣ درجة معيارية يقابلها المئين ١٤ر
، − ٢ درجة معيارية يقابلها المئين ٢ر٢٨
− ١ درجة معيارية يقابلها المئين ١٥ر٨٧
، صفر درجة معيارية يقابلها المئين ٥٠
+ ١ درجة معيارية يقابلها المئين ٨٤ر١٣
، + ٢ درجة معيارية يقابلها المئين ٩٧ر٧٢
+ ٣ درجة معيارية يقابلها المئين ٩٩ر٨٦ .

عيوب الدرجات المعيارية .

١ − يقوم حساب الدرجات المعيارية على أساس حساب توزيع درجات أفراد العينة ، بإيجاد قيمتى المتوسط والانحراف المعيارى . وهذا التوزيع يختلف باختلاف العينة وظروف الإجراء ومواد الاختبار . ولهذا يحسن أن تشتق الدرجات المعيارية بعد تعديل التوزيع التجريبى للدرجات إلى توزيع

اعتدالى حتى يكون أساس الحساب ثابتاً . وتسمى هذه العملية « تسوية المنحنى Normalization » .

٢ – قد تكون إشارة الدرجة المعيارية سالبة ، ولهذا تهدف تعديلات الدرجة المعيارية إلى تسوية المنحنى وإلى التخلص من الإشارات السالبة بجعل كل الدرجات المعيارية موجبة .

٣ – تكون الدرجة المعيارية غالباً كسراً عشريا أو درجة وكسراً عشريا. وعلى هذا يعبر عن الفروق بين الأفراد بوحدات صغيرة جداً لا تمثل مدى هذه الفروق . ولهذا أيضاً هدفت تعديلات الدرجة المعيارية إلى توضيح هذه الفروق والتخلص من الكسور .

نسبة الذكاء الإنحرافية Deviation I Q :

وهى فى الحقيقة درجات معيارية معـدلة ، إلا أن متوسطها الحسابى ١٠٠ وانحرافها المعيارى ١٦ فإن + ١ درجة معيارية تقابل ١١٦ وهى بهذا تتخذ نفس متوسط وانحراف مقياس ستانفورد بينيه وذلك لشيوع استعمالها ودلالة نسب ذكائه ووضوحها . وبهذا فالشخص الذى تكون درجته المعيارية + ٢ فنسبة ذكائه الانحرافية = (+ ٢ × ١٦) + ١٠٠ = ١٣٢ .

وهذه النسب كما هو واضح لا تأتى من قسمة العمر العقلى على الزمنى مع ضرب النتائج فى ١٠٠ كما هو معهود فى نسب الذكاء المستخدمة فى اختبار ستانفور بينيه . ولكنها تأتى من استخدام الدرجات المعيارية بتحويلها إلى توزيع متوسط ١٠٠ وانحرافه ١٦ حتى تسهل مقارنة نسب الذكاء على اختبارات مختلفة .

أنواع أخرى من الدرجات المعيارية المعدلة

١ – الدرجة ت :

الدرجة ت T. Score درجة معيارية معدلة تهدف إلى التخلص من

عيوب الدرجات المعيارية . والاسم ت مأخوذ عن الحرف الأول من اسم ثورنديك Thorndike . والفكرة الأساسية فيها هى تحويل الدرجات المعيارية فى أى اختبار إلى درجات معيارية اعتدالية متوسطها الحسابى ٥٠ والانحراف المعيارى لها ١٠ . وهكذا تتراوح درجات أفراد أى مجموعة بين ـ ٥ ، + ٥ درجة تائية :

فإذا فرضنا أن اختبارا طبق على مجموعة من الأفراد فحصل الفرد س على الدرجة الخام ١٨ . وكان المتوسط الحسابى للمجموعة هو ١٥ والإنحراف المعيارى لها هو ٦ . فتكون درجته المعيارية هى :

$$\frac{١٨ - ١٥}{٦} = ٠٫٥$$

وتكون درجته التائية هى :

$$٠٫٥ \times ١٠ + ٥٠ = ٥٥$$

وإذا كانت الدرجة الخام للفرد ص فى نفس المجموعة هى ٩ تكون درجته المعيارية هى :

$$\frac{٩ - ١٥}{٦} = -١$$

وتكون درجته التائية هى :

$$-١ \times ١٠ + ٥٠ = ٤٠$$

وهكذا نرى أن الدرجة التائية هى درجة معيارية اعتدالية انحرافها المعيارى ١٠ ومتوسطها الحسابى ٥٠ . وهى تحسب من الدرجات المعيارية وليس من الدرجات الخام . ولكن يعيبها أن الفروق بين درجات الأفراد

تمثلها وحدات صغيرة قد لا تفيد فى بعض البحوث العملية التى تكون فيها الفروق بين الأفراد ، بطبيعة الحال صغيرة .

٢ — ولهذا نشأت درجات تائية معدلة فمثلا فى إحدى اختبارات الجيش الأمريكى ، واسمه اختبار التصنيف العام للجيش Army (A.O.C.T) General Classification test، كان الانحراف المعيارى ٢٠ والمتوسط الحسابى ١٠٠ فتكون الدرجة المعيارية لمن يحصل على ١٢ فى اختبار متوسطه ١٠ وانحرافه المعيارى ٤ هى :

$$\frac{١٢ - ١٠}{٤} = ٥ر٠$$

وتكون درجته المعيارية ، بحسب هذا المعيار الحربى ، هى :

$$٥ر٠ \times ٢٠ + ١٠٠ = ١٢٠$$

٣ — كما لجأت بعض الجامعات الأمريكية إلى تعديل الدرجة التائية باستخدام متوسط حسابى ٥٠٠ وانحراف معيارى ١٠٠، حتى تتضح الفروق أكثر . فإذا كانت الدرجة المعيارية لفرد ما هى — ٣ر٠ فإن درجته المعيارية بحسب هذا المعيار الجديد هى :

$$- ٣ر٠ \times ١٠٠ + ٥٠٠ = ٤٧٠$$

٤ — الدرجة جـ :

الدرجة جـ C Score هى تعديل للدرجات المعيارية بحيث تصبح درجات معيارية اعتدالية . اشتقها Guilford من توزيع متوسطه ٥ وانحرافه المعيارى ٢ . وكان هدفه هو إيجاد درجات معيارية اعتدالية تناسب بعض الاختبارات التى تقيس وظائف تكون الفروق بين الأفراد فيها واضحة . كما أن مزايا هذا المعيار قلة عدد المستويات التى ينقسم إليها الأفراد ، وطريقة حسابها هى :

إذا كانت الدرجة المعيارية للفرد هى ٤ر٠

تصبح درجته بالمعيار ج هى :

٤ر٠ × ٢ + ٥ = ٨ر٥

٥ – معيار التسيع :

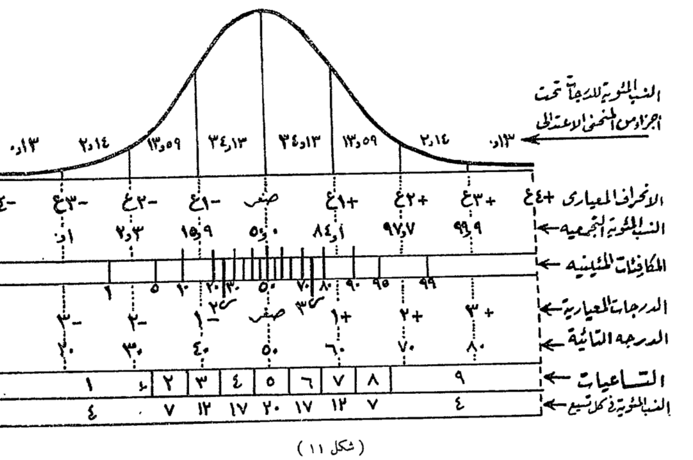
معيار التسيع Standard Nine or Stanine يقوم على أساس الدرجة ج . فبدلا من تقسيم التوزيع إلى ١١ قسما أصبح سلاح الطيران الأمريكى يقسمه إلى ٩ أقسام . وذلك بضم الدرجة ج (صفراً) والدرجة ج (١) معاً فى التسيع (١) وضم الدرجة ج (٩) والدرجة ج (١٠) معا فى التسيع (٩) .

وهكذا يوضح هـذا المعيار الفروق بين الدرجات إذا لم يكن مداها كبيراً ، خاصة إذا كان الأفراد المتطرفون سلبياً متقاربين من بعضهم وكان الأفراد المتطرفون إيجابياً قريبين من بعضهم أيضاً . وطبيعى ألا يصلح هذا المعيار إذا كان الأفراد موزعين فى مستويات تشمل أقساماً دنيا وأخرى عليا .

تحويل المعايير

يجب أن يقدم كل صاحب اختبار لمن يستخدمونه قائمة معايير Table of Norms اختباره موضحاً فى دراسته عدد ونوع ومستوى وطريقة اختيار وتركيب العينة المستخدمة والعمليات الإحصائية وشكل التوزيع . ذلك حتى يتمكن الأخصائيون فى شؤون القياس من مقارنة الدرجات الخام أو المحولة التى يحصل عليها الأفراد فى اختبارات مختلفة فى معاييرها . وهكذا يمكن تحويل درجة الفرد سواء كانت خاماً أو معياراً إلى معيار آخر .

بل ويحسن أن توضح قائمة المعايير « المقابلات المعيارية المختلفة لكل درجة خام » حتى أنه بمجرد النظر إلى القائمة نعرف ما تساويه الدرجة الخام من معيار الفرقة الدراسية أو العمر أو الدرجة المعيارية أو التائية أو . الخ .



(شكل ١١)

المنحنى الاعتدالى مبينا عليه بعض المعايير

والشكل (١١) يوضح المنحنى الاعتدالى مبيناً عليه أنواع مختلفة من المعايير المتقابلة .

النسب

١ ــ أول هــذه النسب Quotients هى نسبة الذكاء I. Q فإذا كان العمر العقلى للفرد هو ٩٠ شهراً وكان عمره الزمنى ٧٥ شهراً كانت :

$$\text{نسبة ذكاوه} = \frac{٩٠}{٧٥} \times ١٠٠ = ١٢٠$$

وإذا كان العمر العقلى للفرد ١١ سنة و٣ شهور والعمر الزمنى له ١٢ سنة و ٦ شهور كانت :

شهر سنة

$$\text{نسبة ذكاوه} = \frac{١١ - ٣}{١٢ - ٦} \times ١٠٠$$

$$\text{وبتحويل السنوات إلى شهور} = \frac{١٣٥}{١٥٠} \times ١٠٠ = ٩٠$$

٢ ــ أما نسبة التعليم (.Educational Quotient (E. Q فهى النسبة بين تحصيل الفرد وعمره الزمنى . ونحن نعلم أن التحصيل يقاس بحساب الفرقة الدراسية التى تتكافأ درجة الفرد مع متوسط درجات تلاميذها فى مادة الاختبار .

فإذا فرضنا أن اختباراً للحساب قنن على عدة فرق دراسية وأن التلميذ س حصل على الدرجة ٨ وأن التلميذ ص حصل على الدرجة ١١ وأن التلميذ ع حصل على الدرجة ١٥ . وكان متوسط درجات السنة الأولى الإعدادية هو ٦ والثانية هو ١٠ والثالثة هو ١٢ والأولى الثانوية هو ١٤ والثانية الثانوية هو ١٥ . فإذا كانت أعمار التلاميذ الثلاثة على الترتيب هى ١٠ سنوات ، ١١ سنة ، ١٢ سنة و ٦ شهور .

فإننا نستطيع أن نعد السنة الأولى الإعدادية هى السنة السابعة والثانية الإعدادية هى الثامنة والثالثة الإعدادية هى التاسعة والأولى الثانوية هى العاشرة

(ما دامت المرحلة الابتدائية تنتهى بالسنة السادسة) وهكذا يكون حساب العمر التحصيلى للتلميذ س بالطريقة التالية :

١٠ – ٦ = ٤ مدى الفئة أو الفرقة السابعة أى الأولى الإعدادية . قطع هو من هذا المدى ٨ – ٦ = ٢ أى أن فرقته هى السابعة مضافاً إليها نسبة ما قطعه من هذا المدى أى $\frac{٢}{٤}$ فتكون فرقته ٥ر٧ . وبفرض أن متوسط عمر هذا المستوى الدراسى هو ٥ر١٢ .

وتكون فرقة التلميذ ص = ٨ + $\frac{١١ - ١٠}{١٢ - ١٠}$ حيث ٨ هى الفرقة الثانية ، ١٠ هى درجته ، (١٢ – ١٠) هى مدى الفرقة الثانية ، (١١ – ١٠) هو المدى الذى قطعه من هذه الفرقة . ونسبة ما قطعه من تحصيل إلى ما يجب أن يقطعه إلى آخر العام ، حتى ينقل إلى الفرقة الثالثة ، هى $\frac{١}{٢}$ أى أن فرقته ٥ر٠٨ ومتوسط عمر هذا المستوى الدراسى هو ٥ر١٣ .

وبالمثل يمكن حساب فرقة التلميذ ع = ١٠ + $\frac{١٥ - ١٥}{١٥ - ١٤}$ = الفرقة ١٠ ومتوسط عمر هذا المستوى الدراسى هو ١٥ . فتكون أعمار الثلاثة التحصيلية على الترتيب هى ٥ر١٢ ، ٥ر١٣ ، ١٥ .

النسبة التعليمية (E.Q) = $\frac{\text{العمر التحصيلى A.A}}{\text{العمر الزمنى C.A}}$

النسبة التعليمية للتلميذ س = $\frac{١٢٫٥}{١٠}$ × ١٠٠ = ١٢٥

النسبة التعليمية للتلميذ ص = $\frac{١٣٫٥}{١١}$ × ١٠٠ = ١٢٣

النسبة التعليمية للتلميذ ع = $\frac{١٥}{١٢٫٥}$ × ١٠٠ = ١٢٠

٣ — النسبة التحصيلية Achievement Quotient (A.Q.) هى النسبة بين تحصيل الفرد وعمره العقلى محسوبا عن اختبار للذكاء . وقد تسمى « نسبة الإنجاز » Accomplishment Q. (A.Q). فإذا كانت نسبة ذكاء الفرد هى ١٢٥ ونسبته التعليمية هى ١٠٠ كان مستواه العقلى يسمح له بنجاح أكثر من الذى حصله وكانت نسبة تحصيله هى :

$$\text{نسبة التحصيل} = \frac{\text{نسبة التعليم}}{\text{نسبة الذكاء}}$$

$$= \frac{١٠٠}{١٢٥} \times ١٠٠ = ٨٠$$

وإذا كان العمر التحصيلى لفرد هو ١٢ سنة وعمره العقلى هو ١٠ سنوات كانت نسبة تحصيله هى :

$$\text{نسبة التحصيل} = \frac{\text{العمر التحصيلى}}{\text{العمر العقلى}} \times ١٠٠$$

$$= \frac{١٢}{١٠} \times ١٠٠ = ١٢٠$$

أى أنه يحصل أكثر مما يسمح له مستواه العقلى وعلى هذا نحدس أنه يبذل مجهوداً أكثر من التلميذ المتوسط . وهكذا تفيدنا هذه المعايير فى معرفة الكثير عن التكوين النفسى والتربوى للفرد .

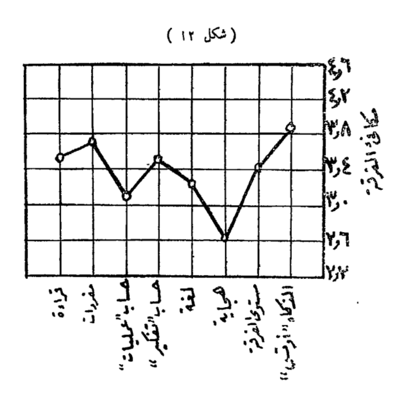
الصفحات النفسية

تمدنا المعايير السابقة بطريقة نعبر خلالها عن درجات مجموعة من الأفراد فى اختبارات مختلفة تماما ، وهذا التعبير يكون بوحدات عامة بحيث تمكن

المقارنة مباشرة . وبالرغم من أن المعايير السابقة تسهل مقارنة درجات الفرد فى اختبارات مختلفة تقيس نفس الشىء ، إلا أن أحــدا من هذه المعايير لا يمكننا من مقارنة درجات الفرد فى اختبارات تقيس وظائف مختلفة .

ولكن إذا حولت درجات أى اختبار يقيس أى وظيفة إلى نوع من المعايير ينسب الفرد إلى مستوى فرقة دراسية معينة فى الحساب مثلا وينسبه اختبار آخر إلى فرقة أعلى فى اللغة العربية وينسبه اختبار ثالث إلى فرقة أقل من الأولى ، سهلت عملية المقارنة . وهكذا يكون للفرد أكثر من فرقة دراسية حسب مستواه فى المواد المختلفة ، إذ أن مستوى أحد تلاميذ الشهادة الإعدادية قد يفوق متوسط درجات تلاميذ السنة الثانية الثانوية وخاصة فى المواد التى تدرس فى الإعدادية ولا تدرس فى الثانوية . انظر الشكل (١٢) وسيتضح هذا الكلام فى قسم القياس التربوى .

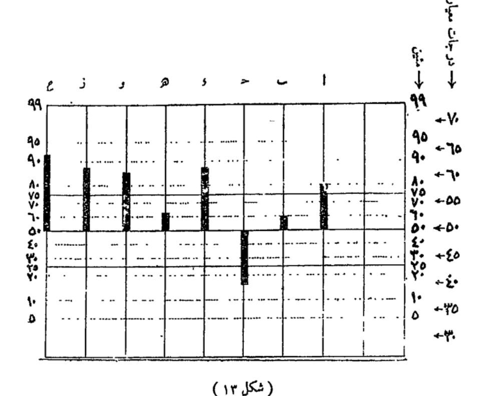
(شكل ١٢)



كما تسهل مقارنة مستويات الأفراد فى وظائف مختلفة إذا عبرنا عن نتيجة كل اختبار فى معيار هو النسبة المئوية لعدد من يفوقهم الفرد من أفراد الجماعة فى الاستدلال اللفظى مثلا ، والنسبة المئوية لعدد من يفوقهم الفرد فى السرعة فى الأعمال الكتابية مثلا . هذا المعيار هو المئين . والمعيار السابق هو معيار الفرقة الدراسية .

وليست هذه المعايير جديدة علينا ، فقد سبق عرضها ، ولكن الجديد هو جمعها فى صفحة واحدة هى الصفحة النفسية Psychoprofile وهى عرض بيانى graphic presentation مجمع لدرجات الفرد فى اختبارات مختلفة . وهناك صفحات تستخدم لتسجيل درجات مجموعة من الأفراد معاً وهى تسمى سجلات Records . فى أحد فئاتها الأسماء وأمام كل اسم سطر ينقسم إلى عدة خلايا كل خلية يوضع بها رقم .

فمثلا فى صفحة سجل الفصل Class Record Sheet فى اختبار كاليفورنيا للنضج العقلى California Test of Mental Maturity ، أمام خلية اسم التلميذ فئة لجنسه وأخرى لعمره الزمنى وفئة ثالثة لعمره العقلى تنقسم إلى عمر عقلى فى الاختبار ككل وعمر عقلى فى القسم اللفظى من الاختبار وعمر عقلى فى القسم غير اللفظى من الاختبار . ثم خلايا لنسبة الذكاء فى الأقسام الثلاث السابقة . وآخر فئة من الخلايا تدرج فيها الرتبة المئينية للفرد فى اختبارات التذكر ، والعلاقات المكانية ، والاستدلال المنطقى ، والاستدلال العددى ، والمفاهيم اللفظية ، وأخيراً خلية للرتبة المئينية للفرد فى مجموعة الاختبارات اللفظية وأخرى لغير اللفظية .



(شكل ١٣)

صفحة نفسية لمفحوص بالفرقة التاسعة على بطارية اختبارات الاستعدادات الفارقية عن D. A. T. Manual, The Psychological Corporation, N.Y.,1947. والحروف تشير إلى الاختبارات وهى :

(ا) استعمال اللغة – الجمل

(ب) استعمال اللغة – الهجاية

(ح) الدقة والسرعة للأعمال الكتابية

(ء) الاستدلال الميكانيكى

(هـ) العلاقات المكانية

(و) التفكير المجرد

(ز) القدرة العددية

(ع) التفكير اللفظى

وفى (الشكل ١٣) صفحة نفسية لمفحوص واحـــد على بطارية اختبارات الاستعدادات الفارقية Differential Aptitude Tests D.A.T. وهذه البطارية تقيس جوانب مختلفة من القدرة . وهذه الجوانب أساس برامج التوجية فى المدرسة الثانوية . ويلاحظ أن الاختبارات المختلفة تمثلها أعمدة منفصلة . والمعايير فيها هى الدرجة المعيارية محسوبة على أساس المتوسط الحسابى والانحراف المعيارى لكل فرقة دراسية على حده . كما أنه ، بالنسبة لكل فرقة دراسية على حده تحسب الرتبة المئينية للفرد على أساس عدد من يفوقهم فى الاختبارات المختلفة .

وواضح أن الاختبارات ، التى تتضمنها البطارية ، هى اختبار لفظى وآخر عددى وثالث للاستدلال ورابع للعلاقات المكانية وخامس للاستدلال الميكانيكى وسادس للأعمال الكتابية (وينقسم بدوره إلى قسمين) وأخيراً اختبار لغوى ينقسم إلى اختبار للهجاية وآخر للجمل . ومن الرسم يتضح لنا أيضاً أنه يمكن التعرف على مقابل الدرجة الخام سواء كان درجة معيارية داخل الفرقة الدراسة التى ينتمى إليها المفحوص ، أم كان الرتبة المئينية للمفحوص بالنسبة لنفس الفرقة . ويمكننا أن نستخرج من الشكل الجدول رقم (٧) :

الاختبار / الدرجة	اللفظى	العددى	المجرد	المكانى	الميكانيكى	الكتابى	الهجاية	الجمل
الخام	٣٣	٢٦	٣٥	٣٩	٣٤	٤٣	٥٤	٤٥
المئينية	٩٢	٨٨	٨٧	٦٣	٨٨	١٧	٦٠	٨٠
المعيارية	٦٥	٦٣	٦٢	٥٤	٦٣	٤١	٥٣	٦٩

الجدول (٧) الدرجات الخام ، للمفحوص فى بطارية اختبارات الاستعدادات الفارقية ومقابلاتها الدرجة المعيارية والرتبة المئينية داخل الفرقة الدراسية

والدرجة المعيارية Standard Scre هنا متوسطها ٥٠ وانحرافها المعيارى ١٠ . وواضح أن توزيع المثينات متركز فى الوسط مشتت فى الأطراف فوحداتها غير متساوية تساوى الدرجات المعيارية المقابلة . ونلاحظ أن توزيع المثينات يقترب من توزيعها فى المنحنى الاعتدالى .

وهذه الصفحة النفسية تبين لنا نواحى القوة والضعف فى المفحوص بالنسبة لمتوسط مستوى جماعته فى كل اختبار على حدة . وهذا المتوسط كان المثين ٥٠ . وعند تفسير الصفحات النفسية يلزم أن نكون حريصين بحيث تمكن المقارنة وتكون صحيحة فتشتق معايير العمر أو الفرقة أو المثين من جماعات متكافئة ، ذلك فى كل الاختبارات . ولضمان تكافؤ كل الجماعات يحسن تطبيق كل الاختبارات على جماعة واحدة .

ولكن يعيب الاعتماد على الصفحات النفسية الفردية أنها تجسم الفروق فى ذات الفرد . ونحن نهتم عادة بالفروق بين الأفراد . وعلى هذا يستخدم هذا النوع من الصفحات فى الخدمات النفسية الفردية : وهناك نوع آخر من الصفحات النفسية كسجل كاليفورنيا يمكن به مقارنة أفراد الجماعة معاً .

اختلاف نتائج الاختبارات

لا تعطى الاختبارات التى تقيس نفس الوظيفة نفس النتائج إذا طبقت عينة واحدة لهذه الأسباب :

١ — اختلاف عينة تقنين كل اختبار عن غيرها والتى على أساسها استخرجت معايير الاختبار . ولهذا يبدو الفرد الواحد متخلفا فى قياس متقدما فى غيره تبعا لاختلاف عينة التقنين فى كل منهما . فإذا كان لدنيا مقياس مقنن على طلبة الجامعات وآخر مقنن على راشدين غير مختارين unselected من نفس السن فأى شخص سيبدو متخلفا فى المقياس الأول عنه فى المقياس الثانى .

٢ ــ اختلاف محتويات المقياس وما يقيسه من وظائف functions وما به من تعليمات وما له من خصائص وفى انحرافه المعيارى ومتوسطه الحسابى عن غيره .

وفيما يلى بيانات تثبت وجهة النظر هذه :

اختيرت ثلاث مجموعات متماثلة فى عمرها وفى فرقتها ومستواها التحصيلى وطبق على كل مجموعة منها واحد من الاختبارات الآتية :

1 — Terman. Mc Nemer test of mental Ability

2— Otis Quick Scoring Mental Ability Test

3 — Pintner general Ability Test verbel Series

ومن نتائج هذه الاختبارات الثلاث عمل جدول لنسب الذكاء المتوازية فى كل الاختبارات الثلاث . وذلك بإيجاد الدرجات التى توازى رتب المئينات فى التوزيعات الثلاثة ، وهذا تحليل لاختلاف الاختبارات الثلاث فى نتائجها :

١ ــ فى اختبار پنتــنر كانت نسب الذكاء أقل من ٢ إلى ٥ عنها فى اختبار ترمان .

٢ ــ كان مقياس اختبار أوتيس ذا انحراف معيارى أقل من كلا مقياس ترمان وپنتنر ، وهذا الاختلاف يجعل نسبة الذكاء قريبة من الوسط فى أوتيس عنه فى المقياسين الآخرين .

٣ ــ كانت نسبة ذكاء ٦٦ فى ترمان تعادل ٧٦ فى أوتيس بينما نسبة ذكاء ١٤١ فى ترمان تعادل ١٣٤ فى أوتيس

ولكى تكون المقارنة بين المقاييس صادقة لن يكفى أن يعبر كل مقياس فى معاييره لوحدات متماثلة (عمر عقلى ، مئين ، درجة معيارية) بل يجب أيضاً أن تكون عينة التقنين متماثلة . وعند ذاك فقط يمكن مقارنة مختلف نتائج المقياس بغيره .

وعلى كل حال فلا بد من نسبة الدرجة إلى المقياس الذى جاءت نتيجته فنقول مثلا إن فلاناً نسبة ذكائه فى مقياس استانفورد بينيه كذا وإن فلاناً نسبة ذكاؤه الانحرافية فى المقياس كذا هى كذا ... الخ .

وذلك لأن معيار المقياس ليس مطلقاً فى دلالته . فقد تعادل نسبة ذكاء ٩٤ على مقياس ما يعادل نسبة ذكاء ١٠٠ على مقياس آخر .

تعليق :

يمكن أن نحسب للاختبار معايير مختلفة ، وذلك لأن الدرجة الخام لا يمكن أن تفهم إلا على أساس هـذه المعايير . ولكل من المعايير المستخدمة ميزاتها وعيوبها ولكن الشائع هو اسـتخدام العمر العقلى ونسبة الذكاء فى اختبارات الأطفال واستخدام المئينات فى اختبارات البالغين .

ومهما كان المعيار المستخدم فيجب أن يفهم على أنه معيار خاص نوعى يتحدد بظروف التقنين وبالعينة التى استخدمت فى التقنين وخارج هذه الظروف وتلك العينة لا يكون من الصواب أن تنسب الدرجة الخام إلى معيار الاختبار .

فالاختبار الذى قنن فى أمريكا يحتاج إلى تقنين جديد واستخراج معايير جديدة إذا أعيد تطبيقه فى أى مجال آخر .

والاختبار الذى يقنن على أنه اختبار للسرعة للكبار قد يكون اختباراً للذكاء إذا استخدم للصغار . معنى هذا أن المعيار له صحة كما أن للاختبار صحة فالمعيار لاختبار السرعة مع الكبار يكون صحيحاً إذا اتخذ على أنه معيار سرعة مع الكبار ولكنه لا يكون صحيحاً إذا اتخذ على أنه معيار السرعة مع الصغار .

كيف نستخدم المعايير

رأينا أن المعايير تقدم لنا أساساً لتفسير درجة الفرد . فتحويل الدرجة الخام عن اختبار على حدة إلى مكافئ عمر أو فرقة دراسية ، أو إلى درجة مئينية أو معيارية ، يسمح لنا بتفسير مستوى قدرة الفرد فيما يقيسه الاختبار المعين . فإذا جعلنا درجات الفرد عن مجموعة من الاختبارات تحول إلى معايير ذات وحدات موحدة ، وعرضناها فى صفحة نفسية ، أمكننا أن نحصل على صورة وصفية للمستوى النسبى لأداء الفرد فى مجالات الفصل ، أو مجموعة الفرقة الدراسية فى المدرسة ، أو أطفال الفرقة الدراسية فى النظام المدرسى كله . إذ بهذا نرى متوسط مستوى أداء المجموعة لوظيفة مفردة ، أو الأداء النسبى للمجموعة فى كل من المجالات المتعددة . فالمعايير تقدم لنا إطاراً يمكن أن ننظر خلاله إلى هذه الصورة ، وأن نرجع كل أجزاء الصورة إلى ذلك الإطار العام . ولكن ما معنى هـــذه الصورة ، وماذا يجب علينا أن نفعل إزاءها ؟

لا نستطيع فى هــذا المجال أن نقدم تفسيراً كاملا لأى مجموعة من الدرجات نحصل عليها من الإجراء فى موقف الاختبار ، وسنقوم بهذا تباعاً فى كل نوع من الاختبارات . إلا أنه بوسعنا أن نضع كل الخطوط العريضة والمبادئ العامة التى تساعدنا فى تجنب الوقوع فى أخطاء تفسير نتائج الاختبار . وسنبدأ بتفسير نتائج الإجراء على جماعة . ثم نتبعه بتفسير درجات الفرد الواحد . ويجب أن ننبه إلى أن التفسيرين يتدخلان فى كثير من النقط .

(ا) مبادئ تراعى عند تفسير درجات مجموعة من الأفراد :

١ – عند تقييم متوسط تحصيل مجموعة من الأفراد ، يجب مراعاة مستوى القدرة المتوسطة لهذه المجموعة . فإذا كان لدينا فصل من فصول

الفرقة الأولى الإعدادية ، وكان متوسط العمر العقلى لتلاميذه هو ١١ سنة ، وجب أن لا نتوقع منه أداء اختبار حساب جودة أداء مجموعة أخرى متوسط عمرها العقلى ١٢ سنة . وإن كان هذا لا ينطبق على كل المواد الدراسية ، فليست العلاقة بين الذكاء والنجاح المدرسى فى كل المواد كبيرة ومتساوية .

٢ - يجب أن نتوقع تأثير نوع ثقافة المفحوص على أدائه للاختبار .

٣ - يجب عند أداء المجموعة أن يتم هذا على ضوء المنهج الدراسى ؛ والنواحى التى أكدت فيه ، وأهداف التعليم العامة والخاصة .

٤ - تستخدم نتائج الاختبارات بطريقة بنائية ، وليس لعقاب المفحوصين .

(ب) مبادئ تراعى عند تفسير أداء الفرد الواحد :

١ - يقيم التحصيل على ضوء معرفتنا باستعداد الفرد .

٢ - يجب أن نضع فى الاعتبار تأثير الأسرة والفروق الثقافية على الأداء

٣ - يقيم أداء الفرد على أساس أنه اشترك فى برنامج دراسى أو برنامج تدريب مهنى معين .

٤ - فى قياس الفرد الواحد يزيد احتمال تعرضنا للوقوع فى أخطاء التجريب .

قائمة مراجع القسم الثانى

1— Adkins, D. C. Construction and Analysis of Written tests for predicting Job performance. Educ. Psy. Measmt., 1946.

2— Adkins, D. C., et al. Construction and Analysis of achievement tests. Washington, D. C. : Government Printing office, 1947.

3— Adkins, D. C. Test Construction in Public Personnel administration Educ. Psy. Measmt., 1944.

4— American Psychological Association. Technical recommendations for Psychological tests and diagnostic techniques. Washington : Amer. Psychol—Assoc. 1954.

5— Anastasi, Anne, and Drake, J. An Imperical Comparison of Certain Tchniques for the Reliability of Speeded Tests. Educ. Psychol. Measmt., 1954.

6— Anastasia, Anne. The Concept of Validity in the interpretation of test scores. Educ. Psych. Measurements 1950.

7— Anastasi, Anne. The Influence of Practice Upon Test Reliability, J. Educ, Psychol, 1934.

8— Anastasi, Anne. Psychological Testing. (2. nd. edition), N.Y. : Macmillan Co., 1957.

9— Anderson, J.E. The Effect of Item Analysis Upon the Diseriminative Power of an Examination. J. Appl. Psychol., 1935.

10— Anonymous. Tables of Norms and Reliability Coefficients. New York : Cooperative Test Service, 1935.

11— Baron, D. and Bernard H. W. Evaluation Techniques for Classroom Teachers. Mc—Graw Hill Book Co. 1958.

12— Bean, K. Construction of Educational and Personnel. Mc — Graw — Hill Co N.Y. 1953.

13— Benett, G K., Seashore, H. G., anp Wesman, A. G. Differential Aptitude Test. N. Y. : Psychol.Corp., 1947.

14— Brown, W. Some Experimental Results in the Correlation of Mental Abilities. B. T. P., 1910.

15— Burt, C., The Reliability of Teachers Assessment of Their Pupils. B. T. Educ. P. Vol. XV, 1945.

16— Clapp, F. L., and Young, P. Self Marking Tests. Boston : Houghton Mifflin, 1929.

17— Conrad, H. S. Characteristics and Uses of Item—Analysis Data. Psychol. Monogr., 1948.

18— Cronbach, L. J. Coefficient Alpha and the Internal Structure of of Tests. Psychometrika, 1951.

19— Cronbach, L. J. Essentials of Psychological Testing. Harper Broths. N. Y., 1949.

20— Cronbach, L. J., and Warrington, W. G. Time — Limit Tests : Estimating their Reliability and Degree of Speeding. Psychometrika, 1951.

21— Cronbach, L. J. On Estimates of Test Reliability. J. of Educ. Psychol., 1943.

22— Cronbach, L. J. Test Reliability : its Meaning and Determination. Psychometrika, 1947.

23— Grow, L. E. and Grow, A. Educational Psychology, American Book Co. (Revised ed) 1958.

24— Curtis, F. D., and Woods, G. G. A Study of A Modified Form of the Multiple — Response Test. J. Educ. Res., 1928.

25— Davis, A. The Measurement of Mental Systems (can intelligence be measured ?) Scient. Mon., 1948.

26— Davis, F. B. Item Analysis in Relation to Educational and Psychological Testing Psychol. Bull., 1952.

27— Davis, F. B. Item Analysis Data. Harvard Educ. Papers, 1946.

28— Davis. F. B. Item Selection Techniques, in Lindquist, E. F., (ed.) Educational Measurement. Washington : Amer. Council Educ., 1951.

29— Dunlap, J. W. Relationship Between Type of Questions and Scoring Errors, J. Exp. Educ., 1938.

30— Flangan J. C. General Considerations in the Selectional Test Items and A Short Method of Estimating the Product—Moment Coefficient From the Tails of the Distribution, J. Educ. Psy. 1939.

31— Flanagan, J. C. Units, Scores and Norms. in E. F. Lindquist (ed.) Educational Measurement. Washington : Amer. Council. Educ., 1951.

32— Freeman, F. S : Theory and Practice of Psychological Testing, New York : Henry Holt and Co. Inc. 1950.

33— Froelich, Clifford P., and Arthur L. Benson : Guidance Testing. Chicago : Science Research Associates Inc. 1948.

34— Garrette, H. E. Statistics in Psychology. (Third edition) Longmans Green and Co. 1948.

35— Ghiselli, E. E., and Brown, C. W. Personnel and Industrial Psychology. N. Y. : Mc Graw—Hill, 1948.

36— Goodenough, Florence L. A. Critical Note on the Use of the Term "Reliability" in Mental Measurement. J. Educ. Psychol., 1936.

37— Goodenough Florence L. Mental Testing, Rinehart, N. Y. 1949.

38— Good, C. V. and Merkel W R. Dictionary of Education. (2 nd edition) Mc Graw—Hill, N. Y. 1959.

39— Greene, E.B. Measurement of Human Behavior. Odessey N. Y. : 1941

40— Greene, L. B. Michigan Vocabulary Profile. J. Higher Educ., 1938.

41— Greene, H. A. and Jorgenson A. N. and Jerberich, J. R. Measurement and Evaluation in The Secondary School. Longmans Green and Co. N. Y. 1947.

42— Guilford, J. P. Psychometric Methods. N. Y. : Mc Graw-Hill, 1936.

43 — Guilford. J. P. New Standards for test evaluation. Educ. Psy. Measmt., 1946.

44 — Guilford, J. P. Fundamental Statistics in Psychology and Education. Mc Graw—Hill Book Co. 1942.

45 — Guilford, J. P. Factor Analysis in a Test development. Psy. Rev., 1948.

46 — Guillihsen, H. The Reliability of Speeded Tests. Psychometrika, 1950.

47 — Guliksen, H., Theory of Mental Tests, 1950. N. Y. : John Wiley & Sons.

48 — Guttman. L. A Basis for Analysing Test — Retest Reliability, Psychom., 1945.

49 — Guttman, L. Prablems of Reliability in Studies in Social Psychology in World war II, 1950.

50 — Holzinger, K. J. Reliability of Single Test Item, J. Educ., 1932.

51 — Horst, P. Estimating Test Reliability from Parts of Unequal length. Educ. P. Meas. 1951.

52 — Humm, D. G. Test Validation on Remote Criteria. J. appl. Psy., 1946.

53 — Jackson, R. W. B. Note on the Relationship Between Internal Consistency and Test - Retest Estimates of the Reliability of a Test. Psychometrike, 1942.

54 — Jackson, R. W. B., and Furgson, G. A. Studies on the Reliability of Test. Univer. Torento, Dep. Educ. Res. Bull., 1941.

55 — Jenkins, J. G. Validity for what? J. Consult. Psy., 1946.

56 — Jones, G., and Galbrich, A. The Interpretation of Standardized tests, Sch. and Soc., 1941.

57 — Kelley, T. L. The Reliability Coefficient. Psychometrika, 1942.

58 — Kelley, T. L., The Selection of the Upper and Lower Groups for the Validation of Test Items. J. Educ. Psy. 1939.

59 — Kuder, G. F, and Richardson, M. W. The Theory of Estimation of Test Reliability. Psychometrika, 1937.

60 — Landis, C., and Kaby, S. E. The Validity of Certain Questions which Purpot to measure Neurotic tendencies. J. Appl. Psy., 1934.

61 — Lowley, D. N. Amethod of Standardizing Group Tests, B. J. Psych. Stat. Sect., 1950. P. P. 86 - 89.

62 — Lawshe, G. H., Jr. A Nomograph For Estimating of Test Items. J. Appl. Psychol., 1942.

63 — Lindquist, E. F., and Cook, W. W., Experimental Procedures in Test Evaluation. J. Exp. Educ, 1933.

64 — Lindquist, E. L. (Ed.) Measurement Handbook. Washington, D. C : American Council on Education, Boston : Houghton Mifflin Co., 1951.

65 — Lorge, I. The Fundamental Nature of Measurement. In. E. F. Lindquist (ed.) Educational. Measurement. Washington : Amer. Council Educ., 1951.

66 — Mc Adory, M. The Construction and Validation of on Art Test. N. Y. : Teachers College Contributions to to Education, 1929.

67 — Mandell, M., and Adkins, D. C. Validity of Written tests for Selection of administrative Personnel Educ. Psy. Measmt., 1946.

68 — Mathews, L. H. An Item Analysis of Measures of teaching ability. J. Educ. Res., 1940.

69 — Mc Call, W. A., How to Measure in Educat.on, 1922.

70 — Mc Clusky, H. Y. An Experimental Comparison of Two Methods of Correcting the Outcome of an Examination. Sch. and Soc., 1934.

71 — Mc Nemar, Q. Psychological Statistics. N. Y. : Willey, 1949.

72 — Meyer, G. The Effect of Recall and Recognition of the Examination Set in Classroom Situation. J. Educ. Psychol., 1936.

73 — Moffie, D. J. Validity of Self — estimated interest. J. appl. Psy., 1942.

74 — Mosier, C. I. A Short Cut in Estimating of Split—Halves Coefficients. Educ.

75 — Mosier, C. I. Factors Influencing the Validity of a Scholastic interest test. J. Educ. Psy., 1937.

76 — Mosier, C. L. A. Critical Examination of the Concepts of face Validity. Educ. Psy. Measmt., 1947.

77 — Mosier, G. I. A Crtical Examination of the Concepts of Face Validity. Educ. Psychol. Measmt, 1947.

78 — Mursell, J. L. Psychological Testing. (2 nd edition) Longmans Green and Co. 1950.

79 — Nunnally Jum C. Tests and Measurements Assessment and Prediction. Mc—Graw—Hill Book Mo. N. Y. 1959.

80 — Otis, S. A. Normal Percentile Chart : Manual of Directions. Yonkers — On - Hudson, N Y. : World Book Co., 1938.

81 — Patterson, C. H. On the Problem of the Criterion in Pridiction Studies. J. Consult. Psy., 1946.

82 — Peatman, J. G. Descriptive and Sampling Statistics. Harper Brothers. 1947.

83 — Recktenwald, L. N. Homageneity of Items in the Cleeton Vocational Interest Inventory. Occupations, 1946.

84 — Remmers, H. H. and Gage, N L. and Rummel, J. F, Practical Introduction to Measurement and Evaluation. Harper Brothers 1960.

85 — Richardson. M. W., and Ruder, G, F., The Calculation of test Reliability Coefficients based upon the Method of Rational Equivalence, V. Educ., 1939.

86 — Richardson, M. W. The Interpretation of Test Validity Coefficient of Increased Efficiency of A Selected Group of Personnel. Psychometrika, 1944.

87 — Richardson, M. W. The Relation Between the Difficulty and the Differential Validity of a Test. Psychometrika, 1936.

88— Ruch, G. M., and Stoddard, G. D. Comparative Reliabilities of Five Types of Objective Examinations. J. Educ. Psychol., 1925.

89— Ruch, G. M., and Stoddard, G. D. Test and Measurement in High School Instruction, N. Y. : World Book Co., 1927.

90— Ruch. F. L. Reporton the Validiation of the Mac Quarrie Test of Mechanical Ability. and the California Short Form Test of Mental Maturity, Language Section, in the Section of Winch Operators. Washington, D. C. : Dept. of Commerce, 1946.

91— Rulon, P. J. A Simplified Procedure for Determining the Reliability of A Test by Split—Halves. Harv. Educ. Rev. 1936.

92— Seder, M. The Reliability and Validity of the A. C. E. Psychological Examination, 1938 Edition J. Educ. Res., 1940.

93— Shaffer, L. F. Information Which Should be Provided by Test Publishers and testing agenciés on the Validity and use of their tests : Personality tests. Proc 1949 Invit Conf testing Service.

94— Sharkey, V. J., and Dunlap, J. W. Study of Reliability and Validity of the Academic Preference Blank. J. Educ. Psy., 1940.

95— Schmidt, H. O. Test Profiles as a Diagnostic aid : The Minnestota Multiphasic Inventory. J. appl. psy., 1945.

96— Schmidt, H. O., and Billingslea F. Y. Test Profiles as a diagnostic Aid : the Bernreuter Inventory. J. abnorm. Soc. Psy., 1945.

97— Schneidler, G. G. Grade and age norms for the Minnesota Vocational Test for Clercal Workers. Educ. Psy. Measmt., 1941.

98— Schultz, R. S. Standardized Tests and statistical Procedures in the Selection of life insurance Sales Personnel. J. appl. Psy. 1936.

99— Skinner, C. E. Educational Psychology. Prentice Hall 1950.

100— Spearman, C. Correlation Calculated from faulty Data. B. T. P. 1910.

101— Staff, Personnel Research Section, A. G. O. Validation of the General Clerical Abilities Test for Selection and Placement of War Dept., Civilian Personnel. Amer Psy., 1947.

102— Staff, Test and Research Section, Bureau of Naval Personell. Psychological test Construction and Research in the Bureau of Naval Personnel : The Basic Test Battery. Psy. Bull., 1945.

103— Stephens, E. W. A Comparison of New England norms with National Norms on the revised Minnesota Paper Form Board Test. Occupations, 1945.

104— Stevens, S. S. On The Theory of Scales of Measurement. Science, 1946.

105— Strong, E. K., JR. Weighted Vs. Unit, Scales. J. Educ. Psy., 1945.

106— Stuit, D. B. The effect of the nature of the criterion upon the Validity of aptitude tests. Educ. Psy. measmt., 1947.

107— Stuit, D. B., and Wilson, J. T. The effect of on increasingly well Defined Criterion on the Prediction of Success at Naval Training School (Tractical Radar). J. appl. Psy., 1946.

108— Super, D. E. The Validity of Standard and Custom-build Personality inventories in a pilot selection Program Educ. Psy. Measmt., 1947.

109— Super, D. E. Appraising Vocational Fitness by means of Psychological Tests. N. Y. : Harper Brothers, 1949.

110— Thomson, G. H. Standardization of group tests and the Scatter of intelligence Quotients B. J. Educ. Psych., 1932, esp, p. 91.

111.— Taylor, K V. F., and Carter, H. D. Retest Consistency of Vocational Interest Patterns. J. Consult. psy., 1942.

112— Taylor, K. V. F. Reliability and Performance of Vocational interests of Adolescents. J. exp. Educ, 1942.

113— Taylor, H. G., and Russell, I. T., The Relationship of Validity Coefficients to the practical Effectiveness of Tests in Selection : Discussion and Tables, J. Ap. 1939.

114— Thorndike R. L. and Hagen. A. Measurement and Evaluation in Psychology and Education. John Willy Sons Inc. N. Y. 1957.

115— Thorndike, R. L. Reliability. in E. F. Lindquist (ed.) Educational Measurement. Washington : Amer. Council Educ, 1951.

116— Thorndike, R. L. Personnel Selection. N. Y. : John Wiley, 1949.

117— Thorndike R. L. (Ed.) Research Problems and Techniques. A A F Aviation Psychology Report, No. 3 Washington, D. C. : Government Printing Office, 1947.

118— Thurstone, L. L. Current Issues in Factor Analysis. Psychol. Bull., 1940.

119— Thurstone. L. L. A Method of Scaling Psychological and Educational Tests, J. Educ. Psychology, 1925.

120— Thurstone, L. L. Scales Construction with Weighted observation, Psy., 1928.

121— Thurstone, L. L., The Absolute Zero in Intelligence Measurement, Psy. Rev. 1928,

122— Thurstone, L. L. The Unit of Measurement in Educational Scales, J. E. Psy., 1927.

123— Thurstone, L. L. Psychological Implications of Factor Analysis, Amer. Psychologist, 1948.

124— Titfin, J. Industrial Psychology. (3 rd ed.) N. Y. : Prentice Hall, 1952.

125 — Trabue, M. R. Grophic Representation of Measured Characteristics of Successful W0rkers. Occupations, 1934.

126 — Travers, R M. W. Educational Measurement. Mac-Millan Co. N. Y. 1955.

127 — Vernon, P. E. Indices of Item Consistancy and Validity. Brit. J. Psychol., Statis. Sect., 1948.

128 — Wadsworth, G. W. Tests Prove their worth to a utality Pers. J., 1935.

129 — Walker, Helen M. Item Selection by Sequential Sampling. Teachers Coll. Rec., 1949.

130 — Warren H. C. Dictionnry of Psychology Houghton Mifflin Co. 1934.

131 — Wilson' J. W., and Carpenter, E. K. The Need for restandadization of altered test. Amer. Psy., 1948.

١٣٢ — السيد محمد خيرى : الإحصاء فى البحوث النفسية والتربوية والاجتماعية . القاهرة ، دار الفكر العربى سنة ١٩٥٦ .

١٣٣ — فؤاد البهى السيد : علم النفس الإحصائى وقياس العقل البشرى . القاهرة . دار الفكر العربى سنة ١٩٥٨.

١٣٤ — محمد عثمان نجاتى : علم النفسى الصناعى . القاهرة ، مكتبة النهضة المصرية سنة ١٩٦٠ .

القسم الثالث

القياس التربوي

الفصل العاشر

التربية والقياس والتقييم

- القياس النفسى والقياس التربوى .
- لماذا تتغير التربية ـ نظرة تاريخية .
- القياس فى التربية .
- تعريف بالتربية .
- معنى التحصيل .
- أهداف التربية .

الفصل العاشر

التربيــــة والقياس

للقياس النفسى والقياس التربوى:

لا يختلف القياس التربوى عن القياس النفسى فى شىء ، بل ولا يختلف القياس فى ميدان عنه فى ميدان آخر فى شىء ، وكل ما نلاحظه هو خصائص نتحدد بها فى ميدان بذاته لأنها تحدد هذا الميدان ذاته . فالقياس ، أين استخدم ، يجب أن يكون موضوعيا مقننا ، والمقياس ، متى أجرى ، يجب أن يكون ثابتا صادقا ، والتقييم ، وكيف أعد ، كيفى كمى يقدر ويفسر ، والدرجة ، لماذا أعطيت ، ذات معنى ودلالة . وأكثر من هذا فالقياس والتقييم يقتصران على موضوع واحد هو سلوك الإنسان باعتباره محصلة عدة قوى تعكس لنا صورة تمثل الإنسان ذكاءه ، وقدراته ، واستعداداته ، وتحصيله ، شخصيته بميوله ، وقيمه ، واتجاهاته ، ورأيه .

وهكذا يشترك القياس والتقييم فى مميزات عامة وخطوط عريضة نبعت أساسا من طبيعة علم القياس وموضوعه الإنسان . ولكن هذا لا يعنى تطابق القياس فى مختلف المجالات والأغراض . فلكل مجال ولكل غرض ، فضلا عن الخطوط العريضة والخصائص العامة المشتركة ، مميزات تخصه . هذه المميزات أتت من الهدف من عمليات القياس فيه وتاريخه وإمكانيات الباحث بالقياس ودوره وكيفية استخدام نتائجه .

إلا أن القياس التربوى والقياس النفسى يقتربان فى الكثير . فإن الغالبية العظمى من الاختبارات والمقاييس جربت وقننت وأجريت على التلاميذ فى مختلف الفرق الدراسية . وعلى هذا فإن هذه الاختبارات تؤتى ثمارها

عندما تستخدم فى المدارس والمؤسسات التربوية . كما أن القياس نفسه نشأ ونمى فى أحضان التربية وما يزال يأخذ من مشكلاتها دفعات قوية يحسن بها أدواته ويوائمها مع المشاكل الحالية . ولا يغيب عنا بالطبع أن دارس التربية درس ، أو يجب أن يدرس ، القياس كما أن المشتغل بالقياس درس التربية أو هو على الأقل قضى جزءاً كبيراً من حياته فى مؤسسات التربية والتعليم أثناء إعداده .

والمشكلات التى نجدها فى المدرسة ، عموما ، هى نفس المشكلات التى تعالجها المقاييس المختلفة فاختبارات الذكاء واختبارات الشخصية أدوات تعين المشتغل بالتربية على المسح ، والتصنيف والإعداد ، والتشخيص والعلاج والتنبؤ و . . الخ عندما يقوم ببحث يدرس فيه حالة تلميذ أو فصل أو النظام التعليمى كله . كما أن الاختبارات التى يضعها المدرس teacher-made وامتحانات النقل والعامة أنواع من المقاييس يمكن بإدخال بعض التعديلات عليها ، كما سنرى فيما بعـــد ، أن تصبح ، أدوات موضوعية تماما كالاختبارات المقننة .

ونظرة سريعة إلى كليات التربية فى بلادنا والخارج نجدها تضم الكثير من أساتذة علم النفس الذين أسهموا بالكثير فى تقدم حركة القياس ، كما أن الكثير من المشتغلين بالقياس كانوا أساتذة فى التربية أو مشتغلين بها . ونظرة إلى كتب القياس التربوى نجدها لا تختلف عن كتب القياس النفسى إلا فى النذر اليسير بل إن الأولى تشمل فصولا من الثانية والثانية تغطى كثيراً من موضوعات الأولى . ونظرة إلى المؤسسات التربوية فى الخارج نجدها الباب الواسع الكبير الذى يفتح للأخصائيين النفسيين بمختلف تخصصاتهم فيفيد منهم ويفيدون منه . ولقد بدأنا نسلك هذا السبيل ، وإن كانت هنالك بعض العقبات ، وإن كان التقدم بطيء ، ولكننا نأمل أن تكون هذه هى البداية ، وأول الغيث قطرة . وكذلك فإن البحوث التربوية إن لم

يقم بها ، فى بلادنا والخارج ، أخصائيون نفسيون ، فإنهم على الأقل يشتركون فيها .

وهكذا تضيق الهوة ، التى قد يتصورها البعض ، بين التربية وعلم النفس أو بين القياس التربوى والقياس النفسى . وهكذا أيضاً نستطيع أن نفيد من الأدوات التى نعرضها هنا فى القسمين التاليين كما أننا سنفيد فى القسم الحالى فعلا من الأدوات التى سنعرضها فيهما على أن يرجع القارئ إلى ما نشير به .

لماذا تتغير التربية ــ نظرة تاريخية

يعتقد رجال التربية أن تقدم العلوم التربوية رهين بالبحوث العلمية التى تهدف إلى اكتشاف الحقائق لتحديد قيمة العمليات والنظم والطرق التربوية ، وأن هذه البحوث يجب أن تتبعها محاولات لتحسين وتطوير نظم التربية والتعليم . ثم توضع هذه النظم الجديدة موضع البحث وتجمع الحقائق العلمية لمعرفة ما إذا كانت تفضل النظم القديمة أم تفتقر إلى بعض مزاياها ؟ وما مدى هذا التحسن ؟ وهل يستحق الجهود التى تبذل فى البحث وفى التغيير أم لا ؟ وهل يمكن أن يطور النظام الجديد أم لا ؟ وإلى أى حد ؟ وما هى التأثيرات المتوقعة لهذا النظام الجديد ؟ وهل يؤدى إلى تغيرات جوهرية فعلا فى إعداد المواطن الصالح وفى تنمية حصيلة التلاميذ أم لا ؟

وهذه الخطوة تستند إلى افتراض مؤداه أن الإنسان كائن منطقى مفكر إذا واجهناه بحقائق تكشف عدم كفاية الطرق التى يتبعها ، فإنه قطعا سيحاول أن يغير هذه الطرق خاصة إذا كان كسبه ، من أى نوع نتيجة التغيير ، كبيراً . ولكن هذا الافتراض ليس صحيحا تماما . فإن تاريخ التربية يدلنا على أن الكثير من النظم التى كشف البحث العلمى عن جدواها لم تفلح فى تغيير بعض النظم التربوية القديمة وخاصة منها ما كان له تاريخه كما أن

تغيير النظم التربوية والتعليمية ، خاصة إذا كان كثيراً ومتقلبا ، مهما كانت فوائده مستقبلا ، يؤدى إلى بلبلة الأفكار واضطراب المتعلمين فضلا عن إعداد جيل جديد من المعلمين يجيد استخدام الطرق الجديدة ويؤمن بالنظام المزمع قيامه فلا فائدة لتشريع لا ينفذ ولا جدوى فى إصدار أمر إلى من لا يستطيع تنفيذه .

يجب ، إذن ، أن نعد المعلمين لهذه التغييرات وندربهم على الطرق الجديدة ونقنعهم بالقيم الحافزة لهذا التغيير . وقد نادى بهذا الرأى أحد رواد التربية فى القرن الماضى وهو چوزيف ماير ريس Joseph Mayer Rice الذى درس الكثير من النظم التربوية فى أوروبا وأمريكا وخرج إلى أن التلميذ كان يتعلم القراءة والكتابة آليا دون أن يكتسب أى أفكار جديدة . والقراءة Reading ، والكتابة Writing ، والحساب Arithmatic (كل منها تبدأ فى الإنجليزية نطقا بحرف R ولهذا تسمى three R's) عند ريس يجب أن نعلمها للطفل بطريقة مشوقة وأن يبدأ تعليمه إياها فى الوقت الذى يسمح له نضجه بأكفأ تعلم ، وأن هذا السن صغير بدرجة . تجعلنا نستطيع أن نعلم الطفل ، ما دامت سنوات تعليمه ستزيد ، الكثير من المهارات التى سيحتاجها فيما بعد .

كانت آراء ريس صحيحة ، ولكن أحدا لم يعره اهتماما . ولكنه سرعان ما عرف أن رأيه لم تؤيده حقائق تعلن للناس ، ومن هنا بدأ يجمع ملاحظاته واستخدم برامج الاختبارات testing programs ليبين أن طرق التعليم المتبعة فى ذلك الوقت غير كافية . وصمم اختبارات فى الحساب ، والهجاية spelling ، واللغات ، ولكنه اكتشف أن جمع الأدلة عن تحقيق أحد الأهداف التربوية عمل صعب ، حتى لو كان الهدف سهلا نسبيا مثل إكساب التلاميذ مهارة الهجاية . فبعد أن أجرى اختبار الهجاية الأول الذى صممه على ١٦٠٠٠ تلميذا ، وجد أن تلاميذ مدارس معينة مستواهم

فى الاختبار منخفض وقد عرف من جولاته أن مرجع هذا إلى عيوب نظام الامتحانات الذى كان يتبع فى هذه المدارس .

ولهذا صمم ريس اختبارا ثانياً للهجاية خاليا من العيوب التى كانت تؤثر فى نتيجة الاختبار الأول ، وهى أن بعض المدرسين ، قلقا فى إعداد تلاميذهم للاختبار ، زودوهم بنطق بعض الكلمات نطقا واضحا ساعدهم فى هجايتها ولهذا اختلفت نتائج هؤلاء التلاميذ عن نتائج غيرهم فى اختبار الهجاية . والنتيجة اللازمة لاستبعاد الكلمات التى يساعد نطقها على هجايتها أن صدق اختباره الجديد ارتفع . وقد انتهى إلى أن مدة تعليم الهجاية فى المدرسة لا تؤثر كثيراً على مهارة التلميذ فى الهجاية إذ أن الأخيرة تتوقف على عوامل كثيرة أخرى ، ولهذا يحسن أن تختصر هذه المدة وأن تستغل فى تعليمه باقى نواحى المنهج الدراسى خاصة وأن هذا الاختصار لن يؤثر على مستواه فى الهجاية كثيراً .

هذا الأسلوب العلمى الذى اتبعه ريس كان صرخة فى وادى مهجور ، فلم تؤثر فى النظام التعليمى فى أى بلد بل إن الكثيرين من مؤرخى التربية يجحدون فضله ما دام لم يغير نظاما رغم علمية طريقته وخطواته التجريبية على أعداد هائلة . والخطأ الوحيد الذى ارتكبه هو اعتقاده بأن عرض الحقائق على رجال التربية سيغير من طرقهم . وقد أفادت البحوث التربوية الحديثة من هذا الخطأ إذ لم تعد تهدف أساسا إلى تطوير البحث العلمى بل أصبحت تعنى أولا بتطوير التربية . ولهذا يجب أن يلم المعلمون بهذه النتائج وأن يتفهمها القائمون بوضع برامج التعليم . ولكن إذا آمنت هيئة تقوم بالإشراف على المناهج الدراسية بأن الوقت قد حان لتغيير نظام تعليمى متبع فإنها بجهودها الضخمة تستطيع أن تقوم ببحوث واسعة تجمع فيها الحقائق وتعرضها بطريقة تضمن التغير إلى أفضل .

ونحن نعتقد أن المدرس يجب أن يحمل عبء المسؤولية الكبرى فى

تقدير نتائج التعليم . وإذا كنا نرى أن الآمال تعقد على المستشارين الذين يوجهون ويخططون هذه الدراسات ، إلا أن الجهد الأكبر يقع على كاهل المدرس نفسه . وهذا التقدير يجب ، كما سنرى ، أن يبعد ، قدر الإمكان ، عن الطرق غير الرسمية وأن يكون أكثر منهجية . ونأمل أن يفيد رجال التربية ، وعلى الأخص المدرسون ، من هذا القسم فى إرشادهم إلى ميادين البحث التربوى وطرقه . ولكن هذا لن يتم إلا إذا اتسعت حركة القياس بما يجعل كل المدرسين مدربين على طرق القياس والتقييم العلمية الموضوعية . وهذا بالطبع لا يعنى أن يصبح كل مدرس خبيرا فى القياس .

القياس فى التربية

فى القياس التربوى نعنى بالقيم التى حققها نظام معين . وهذا النظام قد يكون كل المؤسسات التربوية ، أو برنامجا متضمنا فيها ، أو التخطيط الذى أدى إلى اتباع هذا البرنامج ، أو نتائج هذا البرنامج مثل ما يحصله التلاميذ الذين يعرضون له ، أو . . الخ . وما دام البرنامج التعليمى يقيم خلال ما أحدث من تغيرات فى من اشتركوا فيه ، وهم المدرسون والتلاميذ ، فسنعرض للبرنامج فى فصل ولتقييم المدرسين فى فصل ولقياس التلاميذ فى فصول .

ونتائج البرنامج هى التغيرات الفعلية التى أحدثها فى التلاميذ ، وأهداف البرنامج إذن هى توقع تحصيل هذه التغيرات مستقبلا .

وليس القياس التربوى قاصراً على عملية تحديد النتائج الفعلية التى أحدثها البرنامج الدراسى ، بل إنه يشمل أيضاً إصدار أحكام تتعلق بماهية النتائج المرجو تحصيلها . ولكن القياس إذا تحول من تقدير النتائج خلال السلوك ، إلى معرفة ما إذا كانت هذه النتائج مرغوب فيها أم لا ، يصبح تقييما evaluation . والحق أننا فى نهاية عمليات القياس نصدر مثل هذه

الأحكام التقييمية حتى نحدد معالم الطريق الذى يجب أن تسلكه البرامج الدراســية .

فإذا حددت الأهداف وصممت المقاييس التى تعرفنا بمدى التقدم نحو هذه الأهداف ، أمكننا أن نعين مدى نمو الفرد فيما يتعلق بهذه الأهداف دون أن نصدر أحكاما تقييمية . وفى هذه الأثناء نصدر أحكام التقيم فقط عندما نقوم باختيار الأهداف التى نود تحصيلها . ومنذ أن نختار الأهداف ، بعد أن قيمناها ، لا نلجأ ثانية إلى التقيم ، بل نقيس المركز الحالى للفرد بالنسبة للمدى الذى قطعه من الطريق إليها . ولهذا يحسن أن نحدد استخدام اصطلاح «التقيم» باختيار الأهداف ، واصطلاح « القياس» assessement بعمليات تعين المركز الراهن للفرد . وقد اصطلح على استخدام تعبير « التقيم » فى عمليات القياس التى تتضمن إصدار حكام تقييميه إلى جوار القياس ، واستخدام تعبير « القياس » فى العمليات التى لا تتضمن إلى جوار القياس إصدار مثل هذه الأحكام .

فتحديد الطرق التى نرغب أن يسير عليها التغير الحادث ، فى التلاميذ بعد دخول البرنامج الدراسى ، هو صلب عمليات اختيار الأهداف المنشودة من البرنامج ؛ وهذه العملية تتم قبل تنظيم البرنامج . وتوضع أهداف البرنامج فى صيغة يفهم منها أنها تغيرات مرجوة ، أو نتائج متوقعة جدير بالبرنامج أن يحققها فى التلاميذ .

ويمكن تحديد مدى فاعلية برنامج تعليمى بطريقة أفضل من تعين التغيرات المطلوبة التى أحدثها . ويجب أن يتم هذا على ضوء الظروف التى تم خلالها عمليات التعلم . ولكن شرح ووصف هذه الظروف ، دون معرفة محددة بكمية واتجاه عمليات التعلم ، لا يكفى لتقييم البرنامج ولا لقياس عمليات التعلم الحادثة . ولكن هذه المعرفة المحددة لا يستطيعها المدرس العادى فى فصله . بل إن التقديرات التى يقوم بها الملاحظون الخارجيون

الأقل خبرة من المدرس بالعوامل المتدخلة فى مواقف معينة والتى تؤثر فى التعلم ، قد تكون أفضل من تقديرات المدرسين حتى الخبراء منهم ما دامت تتم فى الفصل : وواجب القياس التربوى هنا أن يعين المدرس على تطوير وسائل الحكم الدقيق وإصدار تنبؤات صحيحة عما سيحدثه التعلم من زيادة فى المعارف بعد أن نكشف للمدرس عن طبيعة عمليات التعلم وظروفها وطبيعة من يقوم بتعليمه .

تعريف بالتربية

نعنى بالتربية هنا كل العمليات التى تهدف إلى تغيير السلوك : وعلى أساس هذا التعريف يتم جزء كبير من التربية خارج المدرسة ويشمل محاولات الجرائد والمجلات فى تغيير الاتجاهات ، ومحاولات الأشرطة السينمائية ، ومحاولات الإذاعة ، والمحاضرات والندوات ، هذا فضلا عن دور الآباء ، ومحاولات رجال الدين ، ورجال التربية الاجتماعية والرياضية ، ومحاولات كل الأفراد والهيئات الذين يبغون تنمية الفرد وإنمائه بإحداث تغيرات مرغوب فيها بحيث يتجه سلوكه اتجاهات معينا . كما أن تعريف التربية بهذا المعنى الشامل يدخل عوامل لا حصر لها كلها تهدف إلى إحداث تغيرات مرغوبة .

وقد حددت وزارة التربية والتعليم فى الإقليم الجنوبى وظائف التربية فى تقرير رسمى بما يأتى :

١ ــ اكتساب التلاميذ القدرة على فهم العلاقات الاجتماعية الصالحة وممارستها :

٢ ــ مساعدة التلميذ على كشف الصالح من استعداداته وقدراته .

٣ ــ تنمية عادات التفكير المنطقى السليم .

٤ ــ تعود النشاط المنتج .

٥ – تحصيل المعلومات وأدوات المعرفة .

٦ – تكوين جسم سليم .

٧ – تكوين اتجاهات عقلية صالحة .

ربط أهداف القياس بأهداف التعلم :

ترتبط عمليات القياس ارتباطا وثيقا بتحديد أهداف برنامج التدريس ، الذى على أساسه نقيس التحصيل . ومرجع الاختلاف الأساسى فى عمليات القياس والتقييم إلى الاختلافات فى تحديد الأهداف ، بل إن بعض المشتغلين بالقياس التربوى لا يضعون اعتبارا لهذه الأهداف . وإذا أردنا أن نستخدم الاختبارات التحصيلية المقننة وجب أن تتضح فى أذهاننا هذه الأهداف . وعلى ذلك يحسن أن يقوم واضع الاختبار بتحديد الأهداف التى يرى أن التعليم فى المادة المقاسه يهدف إليها ، حتى تصبح الدرجات فى مقياسه ذات دلالة على مدى تحصيل التلاميذ لها .

التقييم الرسمى والتقييم غير الرسمى :

نعالج هنا دراسات التقييم والقياس المنهجية الرسمية ، إلا أنه يجب ألا يغيب عنا أن أغلب عمليات التقييم والقياس فى التربية ليست منهجية ولا كذلك رسمية ، فإن كلا من المدرس والتلميذ يصدر أحكامه عن قيمة الجهد الذى بذله التلميذ ، والمواد المدروسة فى الفصل ، وطريقة التدريس والعلاقات الاجتماعية داخل الفصل و . . الخ . وكذلك فإن الآباء وغيرهم من المواطنين خارج المدرسة يقيمون العمليات التربوية تقيما غير رسمى وغير منهجى ، ولكن هذا التقييم يجب أن يشجع لأنه قد يحمل فى طياته إمارات الديموقراطية الصحيحة واشتراك كل المواطنين فى تقييم أمر يعد فى المكانة الأولى بالنسبة لهم ألا وهو التقدم التربوى .

وإن كان من الواجب أن تتم عمليات القياس والتقييم فى التربية على

أساس منهجى رسمى ، إلا أن هذا يتوقف على إعلان نتائج الكثير من البحوث التى تستغرق جهوداً مضنية وتحتاج وقتاً طويلا ، ولهذا فإن التقييم غير الرسمى غير المنهجى يدفع هذه البحوث . فإذا وضعنا فى الاعتبار أن التقدير المنهجى لنتائج برنامج دراسى محدود يتكلف كثيراً إذ أنه لا يتضمن التعريف الدقيق للنتائج المرجوة وحسب ، بل إنه أيضاً غالباً ما يقتضى إنشاء مقاييس جديدة وجمع الوفير من البيانات لتحديد مدى صدق وثبات هذه المقاييس قبل استخدامها .

كما أننا نعلم أن القياس التربوى ميدان حديث نسبياً ، وأن عدد المقاييس التى يمكن استخدامها فيه عن ثقة عدد محدود . فإذا أضفنا إلى هذا حقيقة أن المدرسين فى تقديرهم لمدى نمو تلاميذهم يميلون إلى المبالغة ، اتضحت لنا قيمة إنشاء اختبارات تربوية جديدة رغم صعوبة وضخامة هذا العمل . ومن الجدير بالذكر أن نشير إلى تعليق برتراند رسل Bertrand Russell عن تدريس الهندسة ، إذ يقول إن هذا العلم ظل إلى وقت قريب يدرس منفصلا عن غيره من الميادين الرياضية مع العلم بأن هذا الفصل غير جائز إطلاقاً . وأهمية هذه الإشارة تنحصر فى وجوب رجوع مصمم المقياس التربوى إلى تاريخ التربية كى يستمد منه حقائقاً تعينه على فهم طبيعة المادة .

مشكلة المحكات فى القياس التربوى :

فى الفصل السادس من هذا الكتاب تعرضنا لمشكلة المحكات ، وإذا رجعنا إلى وبستر Webster وجـدناه يعرف المحك بأنه معيار للحكم ؛ قاعدة كان أو اختبارا ، عن طريقة تقييم الحقائق ، والأسس ، والآراء ، والسلوك عند إصدار حكم صحيح يتعلق بها . وفى كل عمليات القياس والتقييم ، يبدأ الباحث بإقرار معيار معين به يحكم على صلاحية الأداة التى سيستخدمها . وفى برامج التعليم يجب أن توضع هـذه المحكات على

أساس نريد به قياس أثر البرنامج . وهذه المحكات تقوم أولا على وصف ما يستطيعه التلاميذ فى فرقة دراسية معينة بالنسبة لوظيفة يقيسها الاختبار الذى سيطبق عليهم .

وعلى هذا فالمحك فى اختبار القراءة مثلا قد يكون وصفا مفصلا لما نتوقع أن تكون عليه قدرات القراءة لدى تلاميذ عمر معين وفرقة دراسية معينة وذكاء معين . هذه القدرات يجب تعريفها بدقة . ويحسن ألا يقتصر المحك على المستوى المتوقع فقط ، بل وأن تراعى فيه أيضاً كل العوامل المؤثرة فى عمليات التعلم فى موقف ندرسه . والمحك الذى يغلب استخدامه فى القياس التربوى هو النجاح المدرسى . وهو محك يعتقد أنه ثابت طوال حياة الفرد ، بل وإنه يشمل أكثر من التحصيل المدرسى . فأحياناً يؤخذ هذا المحك على أنه دليل عل نجاح الفرد أو توافقه فى مجالات أخرى غير الدراسة . وهذا التعميم غير جائز إذ أنه يشمل جوانب لا يقيسها اختبار تحصيلى واحد ، بل ويشمل جوانبا لا يستطاع قياسها بالدقة المنشودة باستخدام أى اختبار موجود .

وعلى هذا يحسن أن يكون المحك قاصرا على منطقة محددة من التحصيل ثبت أنها ترتبط بالموضوع الذى يقيسه الاختبار . فإذا لم تتوفر فى المحك هذه الصفة ، كان غير مجدى . فإذا افترضنا أننا بصدد قياس نجاح التدريس ، وأننا اخترنا محكا عاما هو مدى ما حصله التلاميذ من أهداف هذا التدريس . رأينا أن قياس نمو التلاميذ يعيننا بطريقة غير مباشرة على قياس نجاح تدريسهم . والواقع أن قياسا من هذا النوع لا يقدر كفاءة المدرس ، إذ أن نمو التلاميذ فى المادة المدروسة لا يرجع فقط إلى كفاءة المدرس بل إن ثمة عوامل أخرى كثيرة تتدخل مثل نموهم العقلى والانفعالى والجسمى وأثر خبراتهم خارج المدرسة . ولما كنا نفتقر إلى مقاييس موضوعية لتقدير كفاية المدرس ، فإننا لا نجد بدا من استخدام محك

تقدم التلاميذ ، على أن نضع فى اعتبارنا دائماً أنه محك تقريبى . على أن هناك طرق أخرى لتقييم المدرس سنورد ذكرها فى فصل آخر .

معنى التحصيل

نعنى بالتحصيل حدوث عمليات التعلم التى نرغبها ؛ وما دام التعريف يتضمن كلمة « نرغبها » فإنه بالتالى يتضمن حكما تقيميا بمعنى أن قيمة التحصيل تتوقف على ما إذا كانت عمليات التعلم تسير فى الاتجاه الذى يعتبر أساسيا فى نظر صاحب اختبار التحصيل . ونحن نقصد بالتعلم عمليات تعلم الحقائق والمعلومات بما فى ذلك الاتجاهات الاهتمامات والقيم ما دامت ، ولو جزئيا ، مكتسبة أى متعلمة . كما أن التحصيل يتضمن أيضاً تعلم أنماط السلوك التى تصبح سمات تميز الشخصية . وبرغم اتساع هذا التعريف إلا أننا غالبا ما نقصره على تحصيل التلاميذ أو اكتسابهم لما تهدف إليه المدرسة أو المدرس أو نظام التعليم عموما . إذ أن ثمة دلالات مؤكدة على تأثير المدرس ، فى تنمية جوانب غير عقليـــة لدى التلميذ ، لم تتوفر بعد .

طبيعة التحصيل :

يجب أن نفرق بين الاستعداد Aptitude للتعلم وبين التحصيل ، أو التعلم فعلا . ويقصد دائماً بالاستعداد للتعلم الذكاء كما تقيسه اختبارات الذكاء . وعلى هذا يفترض أنه لو تساوت الفرص أمام كل المتعلمين ، فإن الفروق بين درجاتهم تعد نتيجة اختلاف درجاتهم للتحصيل أى اختلاف نسب ذكائهم .

وتختلف إختبارات التحصيل المدرسى عن اختبارات الاستعداد العام

فى أن النتائج على الاختبار المدرسى ترجع إلى اختلاف درجات المهارات والمعارف الناتجة عن اختلاف استفادة التلاميذ من التدريب الخاص . ويشتد التداخل بين اختبارات الاستعدادات واختبارات التحصيل فى أحد أنواع اختبارات التحصيل المسمى باختبار الاستعداد Readiness Test : والاستعداد بالمعنى الأخير هو القدرات الخاصة ، والمهارات ، والمعارف التى يتطلبها النجاح فى برنامج معين يمر به الفرد أثناء تعلمه :

والتحصيل هو نتيجة تفاعل عوامل ثلاثة هى :

الاستعداد aptitude للتعلم ، والاستعداد readiness or preparedness وفرصة التعلم .

أهداف التربية

إن الاهتمام المتزايد الذى يبديه المدرسون بقياس وتقييم برامجهم بطريقة منهجية يدفعنا إلى إعادة مناقشة المشاكل المتعلقة بالخطوات الرئيسية فى أى عملية قياس . ونقصد خطوات تحديد الأهداف واختيار محكات القياس المتضمنة :

الأهداف الصريحة والأهداف الضمنية :

أهداف البرامج التعليمية غالباً ضمنية implicit بمعنى أن واضعى البرامج يضعون الأهداف فى أذهانهم ولا يشعرون بالحاجة إلى إعلانها سواء أكان الإعلان كتابة أم كان تحديداً لما يقصدونه . ولا يقلل هذا من شأن البرامج أو الأهداف ما دامت الأهداف واضحة فى أذهان القائمين بوضع البرامج الدراسية : ولكن القائم بالقياس أو التقييم قد لا يفسر أهداف البرنامج الذى يقيسه تأثيره تفسيراً يتفق مع ما كان يقصده واضع البرنامج ،

ولهذا يحسن دائماً أن تعلن الأهداف فى فقرات يسهل تفسيرها . وهذا التحديد الصريح يكون إشكالا أمام مخطط البرنامج الدراسى .

القيم المؤثرة فى إختيار الأهداف :

لا يمكننا هنا مناقشة كل العوامل التى يجب أن يضعها منظم البرامج المدرسية موضع الاعتبار عندما يحدد الأهداف التربوية العامة التى يرغب أن يصل إليها التلاميذ بعد دراسة هذا المنهج . وتزيد هذه العوامل عندما نكون بصدد وضع أهداف مرغوبة للتعليم العام ، وتقل بالنسبة للتعليم المهنى أو الفنى . ففى الحالة الأخيرة تتضح الأهداف إذ أن الغرض من التعليم المهنى عادة هو إعداد أفراد يتميزون بخصائص معينة ، وتتعدد الأهداف عندما نكون بصدد التعليم العام ولهذا يحسن أن نحدد الأسس التى بناءاً عليها يُختار هدف ويرفض آخر .

والموازنة بين أسس اختيار الأهداف ترجع ، على وجه العموم ، إلى الفلسفة التى يتبناها أو يتأثر بها واضع البرنامج . وقد يبدأ رجل التربية بافتراض أن القيم التى يبغى أن يحصلها التلاميذ متمثلة فى البرنامج ، وقد يعتقد أن التلميذ سيكتسب هذه القيم إذا نجح فى البرنامج لأن البرنامج يخدم أغراضاً ترتبط بهذه القيم . وقد يفترض أن البرنامج يجب أن يوجه إلى تنمية قيم مطلقة مثل مواجهة الحاجات النفسية للفرد . وفى حالات أخرى يفكر واضع البرنامج الدراسى فى تنمية النواحى الفكرية لدى الفرد وجعله وحدة متكاملة هى صورة للمواطن الصالح فى مجتمع بذاته وفى ظروف بذاتها يمر بها هذا المجتمع . أما واضع البرنامج الذى ينظر إلى الأمر نظرة علمية فإنه يفكر فى إطار دوافع الأفراد وحاجاتهم التى لو أشبعت ووجهت توجيهاً سليما يتفق مع الأساس البيولوجى والظروف الثقافية وعمليات التعلم فإنها تنمى

الفرد أكمل ما يسمح به تكوينه وأكمل ما يورده المجتمع وهذا هو الهدف النهائى للتربية .

وسنعرض فى الفصل القادم لتقييم المدرس الذى يقوم بتنفيذ البرنامج الدراسى وقياس تقدم التلاميذ نحو أهدافه . وطبيعى ، ما دام النصيب الأكبر من الجهد التنفيذى والتقييمى يقوم به المدرس ، أن نطمئن إليه وإلى كل القائمين على النظام التعليمى .

الفصل الحادى عشر

تقييم المــــدرس

- مميزات المدرس الناجح
- مقتضيات التدريس
- أهم ما يجب أن يتضمنه تخطيطك
- بطاقات يقوم بها المدرس قدرته
- مقياس للتقدير الذاتى لقدرات التدريس
- قائمة مراجعة العوامل الهامة فى شخصية المدرس
- بطاقات محلية لتقييم رجال التربية

الفصل الحادي عشر

مميزات المدرس وتقييمها

مميزات المدرس الناجح :

ماذا يلزمك حتى تصبح مدرساً ناجحاً ؟ يرى البعض أن كل ما تحتاجه هو التفوق فى المادة التى تدرسها . بينما يرى البعض الآخر وجوب تحليك بالنظام والدقة . ويصر آخرون على أن بعض الناس ولد كى يصبح مدرساً ؛ وعلى هذا ففى رأيهم ينجح المدرس الذى خلق لهذا العمل . وبين هذه الآراء المتضاربة يحار المرء .

وربما كانت مشاعر الناس فيما يتعلق بالحكم على مميزات المدرس الناجح أعمق وأصدق من مشاعرهم عند الحكم على أى مهنة أخرى . وهذا يرجع بالطبع إلى أن كل الناس يتتلمذون سنوات طويلة من أعمارهم حين قد لايحتاج الفرد منهم إلى المحامى أو الجراح أو المهندس طيلة حياته . ومتوسط عدد المدرسين الذين يتصلون بالتلميذ المتوسط الذى يتخرج من المدرسة الثانوية هو ٦٠ مدرساً ، ومن هنا نلمس مدى كثرة واستمرار الاتصال بين الأفراد والمدرسين كما أن نجاح التلميذ فى المدرسة يرتبط مباشرة بإمكانيات وقدرات مدرسية . وفى الوقت الذى يهتم التلميذ فيه بمميزات المدرس الناجح ، يهتم الآباء بتقدم أبنائهم مدرسيا .

والشخص الذى يعد نفسه كى يكون مدرساً يهتم أكثر من غيره بمقتضيات النجاح فى التدريس . وهو قد يقوم باسترجاع كل المدرسين الذين اتصل بهم ، ثم يحللهم بعناية وخاصة منهم من يعتبره ناجحاً . وهذه طريقة ناجحة ، ولكن نجاحها يتوقف على صدق وواقعية الأسباب

التى من أجلها يعتبر الفرد مدرساً معيناً ناجحاً : ولهذا وجب عند تحليل المدرسين أن يسجل الفرد هذه الأسباب كتابة ، ثم يطلب من أصدقائه تقييم نفس المدرسين كتابة . وهكذا يستطيع أن يقارن أسبابهم بأسبابه :

وفى هذه المقارنة قد يجد الفرد فروقاً كبيرة فى شخصيات المدرسين . فبعضهم هادئ ، وديع ، ودافئ . والبعض الآخر عدوانى . وأياً كانت الفروق بينهم ، فالاستنتاج الذى يجب أن ينتهى إليه الفرد هو أنه لا يوجد نمط واحد تكون عليه شخصية المدرسين الناجحين . كما أنه من الحكمة أن نقول إن طريقة واحدة للتدريس تضمن النجاح لا وجود لها . فالمدرسون الناجحون شخصيات متميزة لهم طرق مختلفة فى التدريس .

ولا يجب أن ينزعج دارس التربية عندما يجد تباينا كبيراً فى خصائص المدرسين الناجحين ؛ ولا كذلك عندما يصرح بعض هؤلاء الناجحين بأنهم لا يتمتعون بكل الخصائص التى تؤكد نجاح أيهم فى التدريس ، فصحيح أنهم يعرفون القدرات المتطلبة ، ولكنهم يترددون فى تأكيد أن أحدا معينا سيكون مدرسا ناجحا جداً بلا شك ، وأن آخرا لن ينجح .

وأول أسباب صعوبة التنبؤ بالنجاح فى التدريس أن المدرس يتعامل مع أكثر الأشياء تعقدا ــ الإنسان التلميذ . والمدرس الناجح يجب أن يكون قادراً على أن يؤثر فى الناس بطريقة تحدث تغيرات مرغوب فيها فى سلوكهم : ويجب أن يكون قادراً على تشجيع التلميذ فى التفكير وفى اتخاذ قرارات ذكية ؛ فيوجه سلوكه بكفاية متزايدة نحو أهداف إيجابية بنائية ديموقراطية : كما أن تعقد وتركيب أوجه التفاعل بين المدرس والتلاميذ يصل إلى حد لا يمكن معه تحليل هذا التفاعل : وقد سبقنا إلى هذا القول هوراس مان Horace Mann سنة ١٨٥٣ عندما قال :

« التدريس أصعب الفنون قاطبة وأعمق العلوم جميعا . وهو فى كماله

المطلق قد يتضمن تعليم معرفة تامة بالوجود كله ، ومعرفة تامة بالطريقة المحددة التي يجب أن يسير عليها كل تطبيق ممكن » .

وفى تخطيط العمل بالتعليم ، بعد تحليل المدرسين ، عليك تكَشُّف الخصائص المرتبطة عادة بالنجاح فى التدريس . ففى أثناء فحصك لكل مدرس ضع فى ذهنك دائما أن كل خاصية فيه متكاملة مع غيرها ؛ ويجب أن يؤدى تحليلك إلى تكوين إطار عام تستطيع خلاله أن تكون خططا معينة طويلة الأجل لمواجهة هذه المقتضيات requirements أفضل مواجهة ممكنة . وواضح أن كمال خطتك سيعتمد على عمق فهمك ووضوح تمييزك أثناء دراسة هذه المقتضيات .

مقتضيات التدريس

قامت مئات الدراسات لتحليل الخصائص الشخصية والمهنية للمدرسين بغرض تعيين السمات التى يحسن أن يتحلى بها المدرسون ، وتخصيص نواحى القصور . وسنفحص نتائج بعض هذه الدراسات . وسيفيدك هذا فى فهم الأساليب الفنية المتعددة التى استخدمت فى جمع البيانات . وبهذه المعرفة قد تكون قادراً مستقبلا على القيام ببحوث مماثلة . عليك إذن أن تدرس هذه النتائج بعناية ، وأن تحلل نفسك على ضوئها ، وبعد ذلك تستطيع أن تنشئ خططا معينة ومفصلة لمواجهة أكبر عدد من هذه المقتضيات أثناء إعدادك لمهنة التدريس . ولن تقدر على مواجهة كل هذه المقتضيات ، ولكن هذا لا يعجزك فبعضها يتطلب خبرة طويلة ، ودراسة ، وجهداً .

١ ــ دراسة عمل المدرسين :

قام شارترز Charters ، ويبلز Waples سنة ١٩٢٩ ببحث مشكلة تحديد المؤهلات Qualifications الواجب توفرها فى المدرس بدراسة شاملة لما كان يقوم به المدرسون وقتذاك . وبعد تبويب الواجبات التى يقوم بها

المدرسون ، درس الباحثان المهارات والقدرات ، والمعارف التى يتطلبها النجاح فى القيام بهذه الواجبات . وخلص الباحثان إلى وضع قائمة بالسمات نقتطف منها ما يلى :

١ — القدرة على التكيف .
٢ — المظهر الشخصى الجذاب .
٣ — اتساع الميول — إلى المجتمع المحلى ، والمهنة ، والتلاميذ .
٤ — العناية — الدقة ، والتحديد ، والكمال .
٥ — الاعتبار ، التقدير ، العطف ، التعاطف، الذوق ، عدم الأنانية .
٦ — التعاون .
٧ — الوثوق والاتساق .
٨ — الحماس — التيقظ ، . . الخ .
٩ — الطلاقة .
١٠ — القوة — الشجاعة ، الحزم ، الاستقلال .
١١ — الحكم الجيد — التبصر ، الاستبصار ، بعد النظر .
١٢ — الصحة .
١٣ — الأمانة .
١٤ — المثابرة — الصبر ، التحمل .
١٥ — القيادة — المبادءة ، الثقة بالنفس .
١٦ — المغناطيسية — الابتهاج ، التودد إليه ، التفاؤل ، خفة الروح ، الاجتماعية ، لطافة الصوت .
١٧ — الدقة والنظافة .
١٨ — تفتح الذهن .
١٩ — الإبتكار — سعة الخيال ، قوة الحجة .
٢٠ — التقدمية .

٢١ ــ دقة المواعيد .

٢٢ ــ التهذيب ــ سلامة الذوق ، التواضع ، الخلق ، البساطة .

٢٣ ــ التحصيل ــ الذكاء وحب الاستطلاع العقلي .

٢٤ ــ ضبط النفس ــ الهدوء ، الوقار ، . . الخ .

٢٥ ــ الاقتصاد .

٢ ــ سؤال التلاميذ :

وهذا أسلوب آخر شاع استخدامه ويكون بسؤال التلاميذ عن الصفات المتوفرة فى المدرسين الذين يعجبون بهم . وقد وجد أن الأولاد والبنات الذين قضو وقتاً طويلا مع مدرس أقدر من غيرهم على إصدار أحكام أكثر صدقا . وقد قام فرانك هارت Frank W. Hart سنة ١٩٣٤ بالتحدث إلى بعض التلاميذ والتلميذات بعد الأسبوع الأول من بدء الدراسة للتعرف على مدى تقبلهم للمدرسين . فسأل تلاميذ ٦٦ مدرسة فى نهاية المرحلة الثانوية من أنماط وأحجام مختلفة عن المدرسين المفضلين لديهم ، وغير المرغوب فيهم ، وأقيمهم عند التلاميذ . طالبا منهم أن يذكروا أعظم المدرسين الذين قابلوهم طوال حياتهم المدرسية . كما طلب منهم أن يذكروا أسباب كل اختيار . وتدل النتائج التى توصل إليها هارت ، وهى مدونة فى الجدول رقم (٨) ، على أن أهم هذه الأسباب عند البنين والبنات هى السمات الشخصية أو الإنسانية فى المدرس .

(الجـــدول ٨)

أسباب الإعجاب بالمدرس ، مرتبة بحسب تكرار ذكرها ، كما قال بها ٣٧٢٥ تلميذاً بالسنة النهائية من المرحلة الثانوية .

أسباب الإعجاب بالمدرس	تكرار ذكرها	رتبتها
متعاون فى العمل المدرسى ، يشرح الدروس والواجبات بوضوح وشمول ، ويستخدم أمثلة فى تدريسه .	١٩٥٠	١
مبتهج ، سعيد ، طيب ، طروب ، مرح ، ابن نكتة .	١٤٢٩	٢
إنسان ، صديق ، رفيق ، « واحد منا » .	١٠٢٤	٣
ميال ومتفهم لتلاميذه .	٩٣٧	٤
يجعل العمل مسلياً ، يخلق رغبة فى العمل ، يجعل عمل الفصل مسراً .	٨٠٥	٥
حازم ، يضبط الفصل ، أوامره محترمة .	٧٥٣	٦
غير متحيز ، لا يميز أحداً ، (لا محاسيب له) .	٦٩٥	٧
غير متبط ، لا يكثر الشكوى ، لا يؤنب كثيراً ، غير ساخر .	٦١٣	٨
يعلمنا الموضوع .	٥٣٨	٩
شخصية ظريفة .	٥٠٤	١٠
صبور ، عطوف ، متعاطف .	٤٨٥	١١
عادل فى درجاته وتقديراته ، عادل فى امتحاناته واختباراته .	٤٧٥	١٢
عادل، وواسع الحيلة فى تعامله مع التلاميذ ، نظامه محكم .	٣٦٦	١٣
يطلب أداء العمل تماماً وبدقة ، ويجعلك تعمل .	٣٦٤	١٤
معتبر لمشاعر تلاميذه فى حضور الفصل ، حكيم ، يجعلك تشعر بالراحة .	٣٦٢	١٥

أسباب الإعجاب بالمدرس	تكرار ذكرها	رتبتها
يعرف الموضوع ويعرف كيف يشرحه :	٣٠٧	١٦
يحترم آراء التلاميذ ، ويدعوهم إلى المناقشة فى الفصل :	٢٦٧	١٧
ليس متعالياً ، لا يدعى معرفة كل شيء :	٢١٦	١٨
واجباته معقولة .	١٩٩	١٩
شخص معقول ، غير جامد أو متزمت :	١٩١	٥ر٢٠
يتعاون فى المشاكل الشخصية للتلاميذ ، بما فى ذلك ما يحدث خارج الفصل .	١٩١	٥ر٢٠
ملابسه جذابه ، مناسبة ، منسقة ، ذوقها سليم .	١٤٦	٢٢
صغير .	١٢١	٢٣
يعمل بخطة جيدة ، ويعلم ماذا يجب على الفصل أن يعمل :	١١٠	٢٤
مهتم بحماس بالتعليم .	١٠٨	٢٥
يعطى التلاميذ فرصة عادلة للتعويض .	٩٧	٢٦
واجبات المنزل عنده معقولة .	٩٦	٢٧
يتعرف على الفروق الفردية فى القدرة .	٨٦	٢٨
صريح .	٧٨	٥ر٢٩
شخصية جذابة ، حسن المنظر .	٧٨	٥ر٢٩
يدرس أكثر من الموضوع .	٧٤	٣١
مهتم بنشاط المدرسة .	٦٨	٣٢
يدخل إلى الموضوع دون مقدمات .	٥٣	٣٣
حديث .	٥٢	٣٤
حلو ولطيف .	٥٠	٥ر٣٥
صوته جميل .	٥٠	٥ر٣٥
ذكى .	٤٢	٣٧
دقيق ومحب للعمل .	٤١	٣٨

أسباب الإعجاب بالمدرس	تكرار ذكرها	رتبتها
مخلص .	٣٦	٣٩
يعرف أكثر من الموضوع .	٣٢	٤٠
نشيط .	٣١	٤١
يستخدم أحكام سليمة .	٢٢	٤٢
مثقف ومهذب .	٢٠	٤٣

Witty, P. A. "The Teacher Who Has Helped Me Most" National Education Assoc. J., 1947

وقد وجد هارت أن ٨٠٪ من التلاميذ دللوا على أن المدرس الذى يعجبون به هو أيضاً أكفأ المدرسين وإن كان ٥٠٠ر من التلاميذ قد قالوا بأن المدرس الذى يعجبهم أكثر (أى أحسن مدرس) هو المدرس الذى يحبونه أقل من غيره .

وقد قام ويتى Witty بتحليل ١٢٠٠٠ خطاباً أرسلها تلاميذ الفرقة الثانية عشرة تقابل تلاميذ التوجيهية من كل أنحاء أمريكا إجابة على السؤال : « المدرس الذى ساعدنى أكثر » . وكانت سمات المدرسين التى تكررت كثيراً فى إجابات التلاميذ (خطاباتهم) ، مرتبة بحسب تكرارها هى :

١ ـ التعاون والاتجاه الدعوقراطى .

٢ ـ العطف .

٣ ـ اعتبار التلاميذ ، والصبر عليهم .

٤ ـ تنوع الميول .

٥ — المظهر العام والطريقة اللطيفة .

٦ — العدل وعدم التحيز .

٧ — المرح .

٨ — الاستعداد الطيب والسلوك المتسق :

٩ — الاهتمام بمشاكل التلاميذ .

١٠ — المرونة .

١١ — استخدام المدح والتقريظ .

١٢ — تفوق وكفاية فى تدريس موضوع معين .

وقد لخص ويتى استنتاجاته بقوله إن « هؤلاء الأولاد والبنات يظهرون شاكرين للمدرسة بالدرجة التى توفر لهم بها الأمن ، والنجاح الفردى ، والخبرة المشتركة ، وفرص التوافق الشخصى والاجتماعى . وهذه بالتحديد هى العوامل التى تضمن تعليما جيداً » .

وقد قام ويتى بثلاث دراسات تخللت أمريكا كلها محاولا أن يتحقق من نتائجه وأن يوسعها . وقد وجد أن السمات الاثنى عشر المذكورة سابقاً كانت تتردد دائما بالرغم من أن ترتيبها كان يختلف باختلاف الأعوام .

(جــدول ٩)

الأسباب التى من أجلها يقل الإعجاب بالمدرس ، مرتبة بحسب تكرار ذكرها ، كما وردت على لسان ٣٧٢٥ تلميذاً بالسنة النهائية فى التعليم الثانوى

أسباب قلة الإعجاب	تكرار ورودها	رتبتها
مُثبط ، كثير الشكوى ، كثير التأنيب ، ساخر ، سريع الغضب	١٧٠٨	١
غير متعاون فى العمل المدرسى ، لا يفسر الدروس والواجبات ، غير واضح ، العمل غير مُخطط	١٠٢٥	٢
متحيز له محاسيب أو تلاميذ مفضلون ، ويهاجم تلاميذ معينين	٨٥٩	٣
متعالى ، متكبر ، متغطرس ، لا يعرفك خارج الفصل	٧٧٥	٤
وضيع ، غير معقول ، جامد ، لا يحتمل ، غير مهذب ، قاسى جداً ، يجعل الحياة تعيسة .	٦٥٢	٥
غير عادل فى درجاته وتقديراته ، غير عادل فى اختباراته وامتحاناته	٦١٤	٦
غير معتبر لمشاعر تلاميذه يهزأ بهم فى حضور الفصل ، يجعل التلاميذ خائفين ، سريع الغضب	٥٥١	٧
غير مهتم بتلاميذه ولا يفهمهم	٤٢٢	٨
واجباته غير معقولة	٣٥٠	٩
لا نظام له ، لا يضبط الفصل ، لا يجبر أحداً على احترامه	٣١٣	١٠
لا يدخل إلى الموضوع مباشرة ، يذكر كثيرا من الشوؤن الشخصية غير المتعلقة بالموضوع ، كثير الكلام	٣٠١	١١
لم نتعلم ما كان يجب	٢٧٥	١٢ر٥
كئيب ، غبى ، غير مسلى	٢٧٥	١٢ر٥

أسباب قلة الإعجاب	تكرار ورودها	رتبتها
رجعى ، عجوز	٢٢٤	١٤
غير عادل فى معاملته مع التلاميذ	٢٠٣	١٥
يعرف الموضوع ولكنه لا يستطيع توضيحه	١٩٣	١٦
لا يحافظ على المستويات ، مهمل فى عمله	١٩٠	١٧
قاسى ، صعب جداً ، لا يترك فرصة للتعويض	١٨٣	١٨
لا يعرف الموضوع	١٧٠	١٩
لا يحترم أحكام التلاميذ وآراءهم	١٣٣	٢٠
كثير التغير ، غير متسق مع نفسه : غير ثابت	١٢٢	٢١
كسول ، غير ميال للتدريس	١١٥	٢٢
ليس صديقا ، ليس رفيقاً	٩٨	٢٣
يبدى تحيزاً للأولاد أو البنات	٩٥	٢٤
ملابسه غير جذابة أو ذوقها ردىء	٩٢	٢٥
ضعيف الشخصية	٨٥	٢٦
غير مخلص	٧٥	٢٧
شخصية غير جذابة	٦٥	٢٨
لا يعترف بفروق فردية بين التلاميذ	٦٤	٢٩
صوته غير سار .	٦٣	٣٠

Hart,F.W. Teachers and Teaching The Macmillan Co. N.Y., 1934.

وفى دراسة أخرى سئل ٣٩٠٥ تلميذ يمثلون المدارس الثانوية فى مقاطعة إنديانا بأمريكا أن يذكروا المميزات المتوفرة فى المدرس المفضل لدى كل . وقد أجاب أكثر من ٩٠ ٪ بما يأتى : القدرة على شرح الدرس بوضوح ، عدالة الدرجات ، الاستعداد لتقديم مساعدة إضافية لمن

يحتاجها ، خفة الروح ، القدرة على مسايرة المدرسين الآخرين . ثم أتت خصائص أخرى تلى السابقة فى الترتيب من مثل : معرفة الموضوع المدروس ، الاستعداد الطيب ، السعادة البادية مع ميل إلى الابتسام كثيراً ، وصوت الكلام السار .

كما علَّم من ٦٣ ٪ إلى ٧٨ ٪ على الخصائص التالية : فهم مشاكل الطلبة ، عدم التحيز ، جاذبية الملابس ، أن يكون أخا حقا داخل وخارج المدرسة ، الاهتمام الشخصى بالتلميذ ، أن يشبه رجل الأعمال فى الفصل وأن يصر على الدقة ، الإصرار على العمل بجد . وأجاب حوالى ٤٤ ٪ من التلاميذ مختارين الفقرات التالية : ساعدنى فى الخارج عندما كنت فى مأزق ، قال كلاما لطيفا عن عملى ، كان حسن المظهر ، لم يحرجنا فى حضور تلاميذ آخرين ، لم يغضب منى أبدا ، ولم يجبرنى على العمل الشاق . بينما اختار ١٧ ٪ من التلاميذ باقى الفقرات التى سبق أن ذكرناها .

سوف تلاحظ فى دراستنا ومقارنتنا لدراسات هارت ، ويتى ؛ ريتسى Richey ؛ فوكس Fox اتفاقا عاما فيما يتعلق بخصائص المدرسين الذين يرغبهم التلاميذ . وقد ترغب فى استعراض هذه القوائم وتحليل نفسك لترى ما إذا كنت تملك أو لا تملك المميزات المذكورة .

٣ – تحليل تواريخ حياة المدرسين المشهورين :

طريقة أخرى لكسب الفهم بخصائص التدريس هى الفهم الواضح لكثير من التواريخ الشخصية والكتب التى ألفت عن حياة بعض المدرسين المشهورين . ولكن هـــذه التواريخ تعنى خاصة بالصفات وبالعوامل الإنسانية وإن كانت هذه الصفات وتلك العوامل عظيمة الدلالة فى حياة التلاميذ والتلميذات .

٤ ـ دراسة وظائف التربية :

المدرس الناجح هو الذى يتمثل أهداف التربية ويكون فعالا فى تحقيقها وهو الذى يقيس الخبرة التى مرَّ بها فى مدرسته بما يجب حسب أهداف التربية حتى يحـــدد إلى أى مدى حصل هو بعض هذه الأهداف .

٥ ـ قوائم المقتضيات لرجال التربية :

قام كثير من رجال التربية أفرادا وجماعات بوضع قوائم لمقتضيات التدريس مبنية على خبراتهم كمدرسين وكملاحظين وكعاملين مع مدرسين آخرين ، وهذه القوائم تصاغ غالباً فى مقاييس للتقدير Rating Scales تستخدم فى تقييم عمل المدرس فى المدرسة وتقييم طلاب كلية التربية . وقد تكون راغباً فى إستعراض مقياس ربما أتى الوقت الذى يقدر به عملك ، ويحسن هنا أن تدرس الخصائص المتضمنة . وبفهم هذه الصفات جيداً ستكون قادراً على تخطيط خبرات خاصة تحاول أن تكتسبها أثناء إعدادك لمهنة التدريس :

قام مجلس كلية التربية بكاليفورنيا California Council on Teacher Education بنشر قائمة لخصائص العاملين بالتدريس حتى تتضح وظائفهم لطلبة كلية التربية ولإعطاء تعليمات لتقييم طلبة كلية التربية ، وهذه هى القائمة :

قائمة كاليفورنيا للقدرة على التدريس

المدرس القادر :

١ ـ قادر على تعليم التلاميذ :

(ا) يستخدم الأسس النفسية فى التعليم .

١ — يستخدم الدفع بطريقة فعالة ومستمرة .

(ا) يتعرف على ميول التلاميذ ويستفيذ منها ، وكذا قدراتهم وحاجاتهم :

(ب) يفيد من خبرات التلاميذ وينظر إلى مواقف الحياة وإلى الاهتمامات التى يثيرها الموضوع .

٢ — يوفر خبرات تعلم متنوعة .

٣ — يستخدم طرقا متنوعة فى التدريس ، مثل المناقشة والعرض ، :: الخ بطريقة فعالة .

٤ — يضع خططا متعاونا مع التلاميذ .

(ب) يفيد من أسس نمو الطفل والنمو فى مواقف التعلم

١ — يوفر أوجه نشاط متفرقة وواجبات لمواجهة حاجات وقدرات التلاميذ .

٢ — يعرف المكانة الصحية (العقلية والجسمية) لتلاميذه ويكيف أوجه النشاط مع حاجاتهم .

(حـ) يحافظ فى الفصل على جو يساعد التعليم ويتميز بروح الاتزان بين الحرية والأمن .

١ — يُبقى على موقف عمل فعّال .

٢ — يساعد التلاميذ إضطراديا على القيام بالقيادة والمسئولية .

٣ — يوفر للتلاميذ فرصا للتعاون وممارسة القيادة فى أوجه نشاط الجماعات الكبيرة والصغيرة .

٤ — يوفر فرصة للتعبير عن التفكير الناقد المستقل مؤكدا على حرية التعبير وتفتح الذهن .

(د) يضع الخطط بطريقة فعالة .

١ – يساعد الطلبة على تحديد الأهداف الجديرة بالتحقيق فى أوجه النشاط الواسعة ، واليومية داخل الفصل ، وأوجه النشاط الخاصة بالفصل .

٢ – ينظم تدريسه جيداً بأن يختار بحكمة خبرات تعليمية ، ومحتوى الدراسة ومواد التعليم .

٣ – يختار ويستخدم مواد كثيرة التنوع للتعليم (مثل الكتب والكتيبات ، والأشرطة السينمائية ، والمجلات العلمية ، والإذاعة ، والصور والسجلات . . الخ .)

٤ – يفيد من المصادر الموجودة فى مكتبة المدرسة ومكتبة المجتمع المحلى .

(هـ) يستخدم عمليات التدريس المتنوعة

١ – يستخدم عمليات التدريس (مثل التقارير الجماعية ، والمناقشة ، والتخطيط مع التلاميذ) المصححة لتحقيق الأهداف المرجوة من التعليم .

٢ – يعتمد بطريقة فعالة على مساهمة الطلبة فى أوجه نشاط الفصل ؛

٣ – ينمى مهارات المذاكرة لدى التلاميذ ؛

٤ – يستثير أوجه النشاط الخلاقة عند التلاميذ ؛

٥ – يساعد الطلبة على تقييم تحصيلهم الخاص ؛

(و) يستخدم العمليات التشخيصية والعلاجية بطريقة فعالة :

١ – يألف الاختبارات التشخيصية العامة فى ميدانه الخاص والميادين القريبة .

٢ – ينشئ ، ويجرى ، ويفسر الإختبارات التشخيصية ؛

٣ – يستخدم العمليات التشخيصية الأخرى الملائمة ؛

٤ – يخطط ويفيد من العمليات العلاجية .

(ز) يستخدم عمليات صالحة لتقييم تحصيل الطلبة .

١ – يستخدم عمليات التقييم غير الرسمية (سجل الوقائع ، استبيان المقابلة الشخصية) لجمع وتفسير البيانات المطلوبة .

٢ – يستخدم الإختبارات التحصيلية المقننة .

(ا) يألف أكثر الإختبارات شيوعا فى ميدانه .

(ب) يختار ، ويجرى ، ويفسر نتائج الإختبارات ويستخدمها فى التخطيط .

٣ – يستخدم اختبارات المدرس .

(ا) ينشئ اختبارات مناسبة بمهارة .

(ب) يفسر النتائج ويستخدمها فى التخطيط

٤ – يحتفظ بسجلات دقيقة وكافية ، مثل ، الحالات المدروسة ، والسجلات المجمعة .

٥ – يعمل تقارير وافية للطلبة والآباء فيما يتعلق بتقدم الطلبة فى نموهم .

(ح) يدير الفصل بطريقة فعالة :

١ – يخطط نظاما جيداً ثابتاً لمعالجة الخامات ، والأجهزة والمواد .

٢ – يستخدم وقته ووقت التلاميذ بفعّاليه .

٣ – يراقب الراحة الجسمية لطلبته فيما يتعلق مثلا بالحرارة ، والإضاءة ، والتهوية ، والجلسة .

٢ – يرشد ويوجه الطلبة بحكمة :

(ا) يستخدم الأسس النفسية العميقة المتعلقة بتقدم ونمو الأطفال فى توجيه الأفراد والجماعات .

١ – يحافظ على الموضوعية عند معالجة السلوك العدوانى أو غير السوى .

٢ – يكون متعاطفا وحساسا فى مشاكل التلاميذ الشخصية والاجتماعية وأيضاً فى حاجاتهم التعليمية .

٣ – يعمل تعديلات فى المنهج الدراسى وفى الشروط الأخرى على ضوء حاجات التلاميذ .

٤ – يضمن علاقة طيبة كافية مع الطلبة بحيث يأتون متطوعين للاستشارة .

(ب) يحافظ على علاقات فعالة مع الآباء .

١ – يفسر حاجات ، وقدرات ، وميول ، ومشاكل التلاميذ لآبائهم .

٢ – يحصل على تعاون الآباء فى مساعدة الطلبة على مشاكلهم .

(حـ) يجمع ويستخدم البيانات الجوهرية فى الاستشارة .

١ – يطبق اختبارات الاستعداد والذكاء .

٢ – يفسر نتائج مثل هذه الاختبارات .

٣ – يستخدم النتائج المتجمعة فى الاستشارة مع تلاميذه .

٤ – يجعل البحث مناسباً للتوجيه .

(د) يستخدم عمليات الاستشارة الملائمة .

(هـ) يحافظ على علاقات مناسبة مع أخصائيي التوجيه ، عارفا دورهم وحدود مهارته وقدرته .

٣ – يساعد التلاميذ على فهم وتقدير الميراث الثقافى :

(ا) ينظم الفصل لحياة ديمقراطية فعالة

(ب) يوجه الأفراد والجماعات إلى التطبيقات الجوهرية فى شؤون الحياة أثناء عمليات التعليم فى الفصل .

١ – يستخدم موضوعات ميادين معينة لإنماء فهم المشاكل الاجتماعية ، والاقتصادية ، والسياسية .

٢ – ينمى لدى التلاميذ فهما لدلالة موضوعات متعددة تدخل فى المواد الدراسية المقررة .

(ح) يرسم على أساس حصيلته من الخبرة ليحقق النمـــو الثقافى للأفراد والجماعات .

٢ – يساعد الطلبة على أن يعرفوا ويطبقوا على حياتهم اليومية الأسس الديموقراطية التى تأصلت عميقاً فى نمونا التاريخى .

٤ – <u>يساهم بفعالية فى أوجه نشاط المدرسة</u> :

(ا) يخطط تعاونيا وسائل تحصيل الأهداف التربوية .

١ – يشارك بفعالية فى مراجعة المنهج الدراسى ويكون قادرا على تقييم التقدم نحو الوصول إلى الأهداف التربوية .

(ا) يحدد الأهداف بوضوح .

(ب) يجمع البيانات بكفاية ويشتق منها نتائج مناسبة .

(ح) يستخدم عمليات علاجية مناسبة .

(د) يبدى مرونة فى تعديل خططه وخطواته حتى تتلاءم مع خطط وخطوات المدرسة كلها .

(ب) يقوم بنصيبه من المسؤولية فى نشاط المدرسة .

١ – ينفذ بفعالية المسؤولية التنفيذية الموكولة إليه .

٢ – يساهم فى تخطيط وتنفيـــذ أوجه النشاط الخارجة عن المنهج الدراسى .

(ح) يبقى على العلاقات الشخصية المتناسقة مع زملائه .

٥ – يساعد على تكوين علاقات طيبة بين المدرسة وبقية المجتمع المحلى :

(ا) يدرب نفسه على المصادر المتوفرة فى المجتمع ويستخدمها فى أوجه النشاط داخل الفصل .

(ب) يحصل على تأييد الآباء لنواحى نشاط المدرسة .

(حـ) يساعد فى تحديد وحل مشاكل المجتمع المحلى :

١ – يساعد فى تحديد مشاكل المجتمع المحلى وفى تنمية وعى التلاميذ والآباء بهذه المشاكل .

٢ – يستعين بالمصادر المتوفرة والملائمة لدى المدرسة فى مواجهة مشاكل المجتمع المحلى .

(د) يشارك فى شؤون المجتمع المحلى ومشروعاته .

(هـ) يلاحظ أخلاقيات المهنة فى مناقشة المشاكل الخاصة بالمدرسة وخاصة مع رجل الشارع .

٦ – يعمل فى مستوى مهنى :

(ا) يقدم أدلة على الأهمية الإجتماعية لمهنة الآباء ، والطلبة ، وغيرهم من أعضاء المهنة .

(ب) يسير بحسب ميثاق أخلاقى مهنى .

(حـ) يساهم فى المهنة بالعضوية فى المنظمات والمشاركة فى أوجه نشاطها .

(د) يتحمل مسؤولية نموه المهنى الخاص بتخطيط برنامج مناسب للتقدم المهنى .

١ – يستمر فى الدراسة المهنية عن طريق البرامج ، والمحاضرات ، والمعاهد ، والقراءة المهنية وغير ذلك من نواحى النشاط .

(هـ) يساعد فى الإشراف على طلبة كليات التربية وفى توجيه وإرشاد المدرسين المبتدئين .

وهناك الكثير من القوائم عن الخصائص التى يجب توفرها فى المدرس وضعها رجال التربية أفرادا وجماعات .

أهم ما يجب أن يتضمنه تخطيطك

أثناء قراءتك للقوائم السابقة ، قد يدهشك عدد مقتضيات التدريس . والمدرسة فى يومنا هذا يلزمها عدد من المدرسين أكبر منه فى الماضى . ويمكنك أن تلمس هذا فى التوسع الكبير فى نظمنا التعلمية . ففى الماضى كانت المدرسة فى الريف تقتصر أساسا على تعليم الأولاد والبنات المهارات الأساسية فى القراءة ، والكتابة ، والحساب . وكانت الأسرة ، والجامع والكنيسة ، والمجتمع المحلى يقدمون أغلب نواحى التعليم الأخرى الأساسية فى حياة المواطن الصالح . وقد كانت الحياة بسيطة نسبياً . أما اليوم فالحياة أكثر تعقدا ، وهذا يرجع على وجه خاص إلى التقدم التكنولوجى الضخم الذى حصّلناه . وبالمثل حدثت تغيرات جوهرية كبرى فى حياتنا الاجتماعية ، والاقتصادية ، والسياسية . ووظيفة المدرسة يجب أن تتغير وأن تتوسع حتى تواجه حاجات شباب اليوم وشباب الغد . وهذه التغيرات تؤدى بدورها إلى زيادة وتوسع وحداثة مقتضيات مهنة التدريس .

فالمدرس اليوم يحتاج إلى فهم الجو الثقافى الذى تتم فيه كل عمليات التربية والتعليم . وهو يحتاج أن يرى جوهرية التصنيع والتحضر ، والتعاونية ، والتخطيط الاقتصادى ، والهوة الثقافية Cultural Lag ، والتفكك الثقافى ، ونمو الإتجاه العلمى نحو المشاكل ، إلى آخر كل هذه الخصائص المميزة لمجتمعنا المعاصر ما دامت تؤثر على حياة الناس . وعليه أن يكون قادراً على فهم المشاكل الحديثة فى ضوء الظروف التاريخية ، وأن يتعرف على الطبيعة الراهنة لكثير من الشؤون التربوية ،

وأن يستخدم مصادر الماضى عندما يواجه مشاكل الحاضر ، وبناء على ذلك فعليه أن يميز بين الحلول التقليدية وبين الحلول التى آلت إلى الزوال . فهو يحتاج إلى فهم العمليات التى بها تتدخل القيم فى توجيه سلوكنا وأحكامنا .

على المدرس أن يبين فهماً أساسياً يتضح خلال اتجاهاته ، ومهاراته ، وعاداته ، اتجاهاً يتسم بالاشتراكية الديموقراطية التعاونية بالطريقة التى بها تؤثر القيم فى ثقافتنا ، وبالطريقة التى بها ينمى كل فرد نفسه على أكمل وجه ، وبالطريقة التى نحل بها مشاكلنا ، وبالطريقة التى ننظم بها أنفسنا بحسب تغير حياتنا المشتركة . يجب أن نكون على قدر كبير من الفهم الناقد لوظيفة المجتمعات المحلية والهيئات الموجهة للنمو الإنسانى ، وأن نفهم وظيفة المدرسة باعتبارها هيئة عن طريقها يبحث المجتمع عن إمداد وتوسع بغرض تقدم الفرد . يجب أن ينظر المدرس إلى المدرسة باعتبارها نظاماً عهد إليه بأكبر مسئولية فى تحديد اتجاه وطابع نحو المجتمع . فهو يتعرف على الطريقة التى يحسن أن يتبعها فى تدريس الطلبة والطالبات على أن يحددها إدراكه للوظيفة الحقيقية التى تؤديها التربية .

ألقى البحث التربوى ، خلال الأعوام العديدة الماضية ، الكثير من الأضواء الكاشفة على السلوك الإنسانى وعمليات التعلم . فلم تعد التربية تلقيناً وتكديساً للمواد فى أذهان التلاميذ حتى ينجحوا فى الامتحانات . ومدرس اليوم يحتاج إلى فهم شامل لطبيعة النمو الفردى فى كل جوانبه المتنوعة حتى يستطيع أن يتعرف بل وأن يسبق الحاجات المتزايدة لتلاميذه وأن يوفر الظروف التى تسمح بأفضل نمو . يجب أن يكون ملماً بالطرق التى يفضلها الناس وبالطرق التى يختلفون عليها : فيفهم كيف يتفق الفرد مع الآخرين وإن كان فريداً متميزاً أثناء نضجه . وهكذا يفهم فهماً شاملا عميقا طبيعة التعلم الإنسانى ، خلال عادات التعلم ، واتجاهات ، ومهارات ، وقدرات

المتعلم ذلك فى ضوء دوره كموجه فعال لعملية التعلم . وعلى هذا يجب أن تتوفر فيه مهارة الخبير بتوجيه تربية الآخرين نحو الأهداف المرجوة .

بطاقات يقوم بها المدرس قدرته

قد يعنيك استعراض عدد من بطاقات التقويم المصممة بحيث يقيم الفرد نفسه بها . ونحن نعرض لك هنا بطاقتين . ويحسن بك أن تضع علامات على الخصائص المتوفرة فيك . ونأمل أن تكون أمينا مع نفسك فيها ، حتى تتضح لك نقاط القوة والضعف فى نفسك .

وقائمة التقدير الذاتى Self Appraisal صممت فى مدرسة التربية التابعة لجامعة أنديانا لتستخدم فى منهج تمهيدى فى التربية . وهى مبنية على قائمة من ١٥ خاصية يجب أن تتوفر فى القائم بالتدريس ، أعدها أساتذة كلية التربية بجامعة ولاية أوهايو Ohio .

وتدلك القائمة التالية على أهمية تحليلك بسمات شخصية تساعدك على التأثير الصحيح المرغوب فيه على تلاميذك . فالتلاميذ بطبيعتهم ميالون إلى المحاكاة ، وعن طريق هذه المحاكاة يمكنك أن تكون فعالا .

وعليك أثناء استخدام البطاقة أن تكون دقيقا مسائلا نفسك حتى تعرف ما أنت عليه من قوة تساعدك فى التدريس وما أنت عليه من ضعف قد يحول دونك والنجاح أحيانا . ولكن هذا لا يعنى أن كل الخصائص يجب أن تتوفر فيك فهذا مستحيل . وعليك أن تتوقع مقدما أن بعض سمات الشخصية غير المرغوب فيها قد تغطى على البعض الآخر المرغوب فيه .

ولكى نساعدك أكثر من هذا فى دراستك للتربية ، نقدم إليك قائمة للمراجعة تتضمن أهم عوامل الشخصية المؤثرة على نجاحك كمدرس . ويمكنك أن تلجأ إلى زملائك ، وخاصة منهم من يعد نفسه لمهنة التدريس فيقومون بتقييمك على نفس الأسس ثم تقارن تقييمك لنفسك بتقيمهم لك فتتضح لك صورة عن نفسك أكثر صدقا .

مقياس التقدير – الذاتى لقدرات التدريس

يحتاج المدرسون لكى ينجحوا ، أن ينمو قدرات فى ١٥ مجال على الأقل متضمنين فى مقياس التقدير – الذاتى لقدرات التدريس . وواضح أن هذه القدرات ترتبط معاً تبادلياً . والخصائص المدونة تحت كل نوع من الفهم تفيد كنقط مرجعية لتسهيل عملية التقييم . فأنت تقيم نفسك على أساس قائمة القدرات بوضع علامة (✓) فى المكان الملائم . كن أميناً وعادلا مع نفسك فى تعرفك على نقاط القوة والضعف الأساسية حتى تقدم لك تقديراتك معلومات صحيحة عن مؤهلاتك الحالية .

القـــــــــــدرة	دنيا	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١ – « تفهم فرص الخدمة وتقدر على الخدمة »					
١ – يعتمد عليك : دقيق ؛ مستعد لتقبل المسؤولية ؛ تقوم بالعمل عادة حتى تكمله بنجاح .					
٢ – ميال إلى الناس : تستمتع بتواجدك معهم ؛ وتتعاون فى النشاط الجماعى ؛ تحب الأطفال وتعمل معهم بفعالية .					
٣ – مبادئ : غالبا تقوم بالقيادة .					
٤ – الحساسية الإجتماعية : تضع المصلحة العامة فوق الرغبة الشخصية ؛ عارف بمشاكل وحاجات الآخرين .					
٥ – المثابرة : لا يسهل تثبيطك أو إغرائك بالتخلى .					
٦ – التقبل الإجتماعى : متقبل للناس من كلا الجنسين ، والمسنين والشباب ؛ والآخرين من سنك .					
٧ – شخصية متزنة : لست متزمتا للعلم أو للتهريج أو للإجتماعيات ولكن متزنا فى الصفات الطيبة .					
٨ – السلوك : متزن ؛ ومتحرر من القيود والتكلف .					
٩ – المظهر : نظيف ، ومنظم ، وحسن الملبس .					
١٠ – مستعد للتعاون مع الآخرين .					

القـــــــــــدرة	ضعيفا	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١١ – مستعد وراغب فى العمل مع من يختلفون عنك فى الإطار الاجتماعى ، أو الاقتصادى ، أو الجنسى .					
٢ – « تفهم طبيعة تخطيط واستخدام العمليات فى الحياة اليومية والتعليم »					
١ – تعبر عن القيم بوضوح .					
٢ – تصيغ الأهداف ؛ وتعرف ما تريد أن تبحث عنه .					
٣ – ترى الهوة بين ما تريده وما تجده .					
٤ – تنفذ وتحقق الفروض أو البرامج الإجرائية لتحصل على ما تريد .					
٥ – تستطلع المصادر التى عليك أن تعالجها .					
٦ – تختار الفروض والمصادر .					
٧ – تنظم الفروض ، والمصادر ، وأوجه النشاط الإجرائى لحل المشكلة .					
٨ – تتنبأ بالمشاكل والصعوبات التى تقف فى طريقك .					
٩ – ترى الحلول البديلة خلال قيامك بحل المشكلة .					
١٠ – تحس بالعلاقة بين كل جانب من الجوانب سابقة الذكر والجانب الآخر .					
٣ – « تفهم طبيعة تقييم واستخدام العمليات فى الحياة اليومية وفى التدريس »					
١ – توضح معنى الاشتراكية الديموقراطية التعاونية .					
٢ – تصيغ الأهداف واضحة ؛ وتفكر بعناية فى أهداف الحياة ؛ وتضع اعتبارا للعقائد والقيم .					
٣ – تتعرف على المشاكل وتستخدم العمليات التخطيطية فى حلها .					
٤ – « تفهم أسس الحياة الصحيحة ، عقليا وجسميا ، وتطبق هذه الأسس على حياتك الشخصية وعلى التدريس »					
١ – متزن انفعاليا ومتحرر من العاهات الحسية .					
٢ – لك ميول ترويحية عقلية وجسمية معا .					
٣ – لك ميول كثيرة واتصالات شخصية مع أناس متنوعين ، شباب ومسنين ، ذكورا وإناثا .					
٤ – ماهر فى كثير من الميول الترويحية .					
٥ – تستمتع بالحياة .					

القـــــــــدرة	دنيا	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
٥ – « تفهم نمو الطفل وتعلمه وقدرته على تطبيق أسسها على التدريس »					
١ – لديك حب أصيل للأطفال والشباب ورغبة حقيقية فى العمل معهم .					
٢ – كانت لك اتصالات كثيرة مع الأولاد والبنات .					
٣ – لك معرفة علمية عريضة بعمليات التعلم .					
٤ – لك معرفة كافية بالمشاكل الإنفعالية الأساسية ، والجسمية ، والمشاكل الأخرى الخاصة بنمو الأطفال					
٦ – « تفهم الفروق الفردية والقدرة على التعامل الفعال مع هذه الفروق فى التدريس »					
١ – قادر على التعرف على الكثير من الطرق التى يختلف الناس فيها وقادر على التعامل الفعال مع الأنواع المختلفة .					
٢ – تفهم الطرق الفنية فى التعامل مع الفروق الفردية وقادر على تطبيقها فى التدريس .					
٧ – « تفهم عمليات الإرشاد الوظيفى وقادر على تطبيق هذه الأسس فى التدريس »					
١ – تفهم الطرق المتبعة فى التوجيه ، والاختيار ، والإرشاد .					
٢ – تفهم أهمية التوجيه فى المدارس .					
٣ – تعرف متى وكيف توفر التوجيه للآخرين .					
٨ – « تفهم مواد وطرق التعليم والقدرة على استخدامها »					
١ – تفهم سيكولوجية التعلم وفائدتها فى التدريس .					
٢ – تفهم ولديك المهارة فى طرق التعليم .					
٩ – « تفهم مكانة أوجه النشاط الخارجة عن المنهج الدراسى فى المدرسة مع القدرة على توجيهها »					
١ – كانت لك مشاركة فى نواحى النشاط الخارجة عن المنهج المدرسى فى المدرسة الثانوية					
٢ – المشاركة فى كثير من أوجه النشاط الخارجة عن المنهج الدراسى فى الكلية .					
٣ – فهم قيمة أوجه النشاط الخارجة عن المنهج الدراسى .					
٤ – القدرة على توجيه واحد أو أكثر من أوجه النشاط الخارجة عن المنهج الدراسى .					

القـــــــــــدرة	دنيا	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١٠ – « تفهم مكانة التقييم والقدرة على استخدام عمليات التقييم » .					
١ – ماهر فى تفسير نتائج اختبارات التحصيل ، والميول ، والاستعدادات ، والاختبارات العقلية والأنواع الأخرى من الاختبارات .					
٢ – القدرة على تقييم نمو الشخصية على ضوء الاتجـاهات وتغير الطابع .					
١١ – « تفهم المجتمع المحلى والقدرة على ربط المدرسة بالمجتمع المحلى فى التدريس » .					
١ – تفهم أجواء المجتمع المحلى ، وميول ، وحاجات التلاميذ .					
٢ – لديك المهارة فى الطرق الفنية لزيادة فهمك لمصادر المجتمع المحلى وحاجاته .					
٣ – المشاركة الإيجابية فى أوجه نشاط المجتمع المحلى .					
٤ – عضو من أعضاء المدرسة ومهنة التدريس فعال فى المجتمع المحلى .					
١٢ – « تفهم مكانة الوسائط المختلفة فى التعبير الشفوى والكتابى والقدرة على استخدامها فى التدريس »					
١ – تلائم الأفكار مع الموقف ؛ وتعبر عن نفسه بطريقة مهذبة ، ورقيقة ، ومختصرة ، وكافية ، ومنطقية .					
٢ – تعدد العرض بحسب الموقف ؛ وتستخدم طريقة مناسبة للتوافق مع مناسبة الحديث .					
٣ – تنظم أفكارك وتعبر عنها بوضوح وفعاليه ؛ وتتكلم أو يكتب بأسلوب طليق ، ، سا , وبطريقة موحدة .					
٤ – تستخدم قواعد اللغة المقبولة وبناء الجمله فى الكلام والكتابة					
٥ – تخرج الكلمات بطريقة صحيحة ومتميزة .					
٦ – تستخدم صوتك بفعالية ؛ وله مدى كافى ، متحرر من الملل ، وله نغمة ساره ، ورنين وقوة كافيين .					
٧ – يعرض الأفكار كتابة بتنسيق ، ووضوح ، وتنظيم منطقى مقبول .					
٨ – صحيح الهجاية .					

القـــــــــدرة	دنيا	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١٣ – « تفهم مكانة الوسائط الفنية المختلفة فى التدريس والقدرة على استخدامها »					
١ – القدرة على استخدام وإنشاء رسوم بسيطة ، وأشكال بيانية ، ورسومات للأغراض التوضيحية .					
٢ – تفهم الفنون الجميلة والتطبيقية باعتبارها وسائط تدريس فعالة					
١٤ – « تفهم مكانة التحصيل فى العملية التعليمية والقدرة على جعله وظيفيا فى التدريس »					
١ – متمكن من ميدان تخصصه .					
٢ – له معرفة واسعة بميادين المعرفة العامة القريبة .					
٣ – على فهم بالأحداث الجارية فى العالم .					
٤ – ترتيبه متوسط أو متقدم فى المدرسة الثانوية وفى الكلية .					
٥ – نمى اهتماما حقيقيا بالتحصيل – وليس فقط بالدرجات .					
١٥ – « تفهم فلسفة التربية والقدرة على جعلها وظيفية فى التدريس» .					
١ – الاعتقاد فى إمكانية التحسن عن طريق التربية .					
٢ – توجهه فلسفة تربوية تهدف إلى إعداد التلاميذ حتى يصبحوا مواطنين صالحين .					
٣ – تمارس الاشتراكية الديموقراطية التعاونية متسقا مع نفسه .					

Based on a list of 15 competencies for teaching in Student Planning in College, College of Education· Ohio State University, Coumbus, Ohio, 1940.

قائمة مراجعة للعوامل الهامة * شخصية المدرس

تعليمات :

اعتبر كلا من الأقسام العشرة فى هذه القائمة منفصلة . واقرأ كل عبارة تحت العنوان الكبير بعناية وعلم على أى صفة أو سمة تفتقر إليها بوضوح الشخصية المقاسة . وقبل أن تنتقل إلى القسم الثانى راجع أى الفقرات التى خطط تحتها ثم ضع علامة (✓) فى العمود الذى يعبر عن رأيك العام فى الشخص فيما يتعلق بالجانب المقاس من الشخصية .

المستويات المقترحة للسمة أو الصفة	أقل من المتوسط	متوسطة	جيدة	ممتازة
١ – الاستقرار الانفعالى والصحة النفسية :				
متحرر من المخاوف ، والمزعجات ، ومن مشاعر الذل ، والقلق عن الأشياء التافهة ؛ ومستعد لعمل تقديرات واقعية عن إمكانياته العقلية ؛ وغير حساس جدا للنقد ؛ وذا عقل غنى بالإمكانيات الكلية فى الرضى عن النفس ؛ لا يسهل تهييجه ؛ متحرر من الخجل الزائد ومن اللزمات المزاجية وأحلام اليقظة ؛ يحسن مواجهة المواقف غير المتوقعة ؛ يتكيف عن استعداد مع المواقف المتغيرة ؛ يمارس ضبط النفس ؛ متحرر من عقد الدونية والتفوق ؛ له ضبط على حالته الانفعالية ، لا ينقلب فجأة من حالة متطرفة إلى أخرى ؛ يواجه خيبة الأمل فى الحياة بشجاعة تامة .				
٢ – المظهر الشخصى :				
ملابسة تلائم المناسبة ؛ متيقظ ومتزن ؛ يدل مظهره على الاتزان النفسى ؛ له ذوق جيد فى اختيار ملابسه ؛ يعطى انطباعا بأنه مهذب ومثقف ؛ تركيبات الألوان متناسقة ؛ كى الملابس وتنظيفها ؛ تبدو عليه صحة جيدة .				

* جزء مقتبس مع التعديل من :

Raleigh Schorling and Howard T. Batchelder, Students Teaching in Secondary Schools, Mc Graw—Hill Book Co. Inc., N.Y., 1956.

لم نرد عرض الكثير من الاستمارات من هـــذا النوع ، إذ أنها غالباً لا تناسب أغراض التقيم فى بلادنا ، وإن كانت مفصلة شاملة قامت عليها دراسات أثبتت صلاحيتها . ومن الجدير بالذكر أن جهوداً من هذا النوع تقوم الآن فى بلادنا . وإن كان بعضها لا يزال تحت التجربة . وطبيعى أننا لن نعرض بطاقات التقيم هذه كاملة وعلى من يرغب فى الاستزادة أن يرجع إلى ناشريها المدونة أسماؤهم فى آخر هذا القسم .

بطاقات محلية لتقيم رجال التربية

لا يتسع هذا المقام لعرض فقرات هذه البطاقات ، وتناول تعليماتها وطرق إجرائها وتقديرها . وسنكتفى بتمثيل بعضها .

بطاقة تقويم المدرس

قام الأستاذ الدكتور أحمد زكى صالح بكلية التربية بإعداد بطاقة تقويم المدرس وكراسة تعليماتها سنة ١٩٥٩ .

خطوات بناء البطاقة :

١ — تحليل العمل إذ طلب من حوالى سبعين أخصائياً فى التربية والتعليم ، أمضوا خمس عشرة سنة على الأقل فى ممارسة هذه المهنة ، وصف العمل الذى يقوم به المدرس ثم أعدت بطاقة لوصف العمل .

٢ — وبعد حصر الصفات اللازمة للنجاح فى عمل المدرس ، عُرِّفت الصفات المختلفة بطريقة إجرائية ، كما وردت فى تقارير وصف عمل المدرس ثم عرضت هذه الصفات على مجموعة من الحكام المختصين فى التربية والتعلم ، يبلغ عددهم حوالى ٨٠ مقدراً ، وطلب منهم تقدير درجة لكل صفة من هذه الصفات . وكان التقدير حسب أهمية الصفة فى عمل المدرس ، يعطى فى صورة رقمية من عشر درجات . وقد استبعدت أى صفة كان متوسطهما

أقل من خمسة . وهكذا بلغ عدد الصفات خمس وعشرون صفة . أضيفت إليها خمس عشرة صفة أخرى تبين أن البطاقة بوصفها الأول كانت تنقصها . وهكذا بلغ مجموع الصفات أربعين صفة .

٣ ــ تحديد الصفات : جمعت الصفات ووجد أنها تتميز إلى مجموعتين كل منها عشرين صفة ، المجموعة الأولى هي الصفات الشخصية والثانية هي الصفات المهنية . ووضعت أمام كل صفة خمس درجات تكون سلماً مصاغة في أسلوب وصفي يمكن تحويله إلى أرقام . واستخرجت الدرجات المعيارية بطريقة ليكرت Likert .

وإليك بعض فقرات هذه البطاقة :

الصفة	الدرجة	الصفة	الدرجة
١ ــ القدرة على اتباع التعليمات		٥ ــ الدقة في إعداد الدروس	
٢ ــ العلاقة بالرؤساء		٦ ــ الدقة في تعبيراته الشفوية	
٣ ــ الاتزان الانفعالي		٧ ــ الدقة في الأعمال التحريرية	
٤ ــ إدراكه للبيئة المحلية		٨ ــ اتجاهه إزاء الآراء الحديثة في التربية	
مجموعة الدرجات في الصفات الشخصية		مجموعة الدرجات في الصفات المهنية	

المجموع الكلي :   [   ]   التوقيع

وإليك طريقة تقدير بعض هذه الصفات :

الصفات المهنية

٢١ ــ الدقة في إعداد الدروس :

وتتمثل في عناية المدرس بإعداد دروسه سواء في كراسة التحضير أو في

كراسة خارجية ، بحيث يضمن جميع احتمالات أسئلة الطلاب مع مراعاة الرجوع إلى مراجع غير الكتاب المقرر ، وشرح الطريقة التى تتبع .

٥ ــ ممتاز　　　٤ ــ جيد

٣ ــ متوسط　　　٢ ــ أقل من المتوسط

١ ــ مهمل .

٢٢ ــ الدقة فى تعبيراته الشفوية :

وتتمثل فى اختيار الألفاظ التى تعبر عن أفكاره فى دروسه ومناقشاته .

٥ ــ دقيق جداً فى تعبيراته الشفوية .

٤ ــ واضح فى تعبيراته الشفوية .

٣ ــ يحتاج إلى شىء من التوجيه فى تعبيره الشفوى لكى يحسنه .

٢ ــ لا يحسن التعبير الشفوى .

١ ــ يخطئ فى تعبيراته الشفوية خطأ كبيراً .

٢٣ ــ الدقة فى الأعمال التحريرية :

وتتمثل فى الدقة فى التعبير اللغوى بقصد تحديد المعنى بطريقة واضحة مختصرة مع تجنب الأخطاء اللغوية .

٥ ــ دقيق فهماً ولغة .

٤ ــ دقيق فهماً وليس كامل لغويا .

٣ ــ ينضج فى الأعمال التحريرية بعد التوجيه .

٢ ــ ينقص المعنى .

١ ــ كثير الخطأ فى الأعمال التحريرية .

جهود وزارة التربية والتعليم

قام فرع التقويم بإدارة البحوث الفنية بوزارة التربية والتعليم بالإقليم الجنوبى عامى ١٩٥٨ ، ١٩٥٩ بتصميم :

١ – دليل تقويم المدرس .
٢ – دليل تقويم المدرس الأول .
٣ – دليل تقويم الناظر – الوكيل .
٤ – دليل تقويم مفتش القسم .
٥ – دليل تقويم العضو الفنى .
٦ – دليل تقويم الأخصائى الاجتماعى .
٧ – دليل تقويم الموظف الإدارى .
٨ – دليل تقويم الموظف الكتابى .
٩ – دليل تقويم الرئيس الإدارى .
١٠ – دليل تقويم رئيس القسم أو الوكيل .

وقد أشرف الدكتور محمد خليفة بركات على إعدادها . وقام بالإعداد السيد محمد البسيونى الخطيب . واشترك فى التجريب السيد رجاء محمود علام . واشترك فى مناقشتها الخبراء من العاملين بكل ميدان منها . وكلها تحت التجربة . وهى ترسل بالبريد لعينات من الفئات العشر الجارى تقويمها .

وإليك بعض فقرات من دليل تقويم المدرس :

موضوع التقدير :	مراتب التقدير	الدرجة
أولا : الإلمام بالمادة والاستزادة فيها مدى إلمامه بمقررات المواد التى يقوم بتدريسها واستزادته فيها .	(ا) يتقن مادته إتقانا تاما ويحرص على الاستزادة منها باستمرار سواء بالاطلاع أو الاشتراك الفعال فى الندوات والمؤتمرات	١٢
	(ب) له إلمام كبير بالمواد التى يقوم بتدريسها ، ويطلع على كثير مما يجد منها ، ويذهب أحياناً إلى الندوات العلمية	٩
	(ج) يلم بالمقررات التى يدرسها ، ويطلع أحيانا على ما يجد منها	٦
	(د) تقتصر معلوماته على الكتب المقررة ، وقلما يطلع على الحديث من مادته	٣

موضوع التقدير :	مراتب التقدير	الدرجة
ثانياً : الإلمام بالعملية التربوية وطريقة التدريس : مدى إلمامه وفهمه لأهداف تربية التلميذ وانعكاس ذلك على سلامة اختياره وتطبيقه لطرق التدريس المناسبة لأحوال التلاميذ .	(ا) يفهم تماما أغراض العملية التربوية فى المجتمع ويجيد اختيار طرق التدريس بما يلائم مستويات التلاميذ وظروف بيئتهم	٨
	(ب) يدرك أهداف العملية التربوية ، ويحاول أحيانا تطبيق طرق التدريس المناسبة	٦
	(ج) معلوماته عن العملية التربوية غير وافية ، وقليل التجديد فى طرق تدريسه	٤
	(د) معلوماته سطحية عن العملية التربوية وأهدافها ، ويخلط دون تبصر بين طرق التدريس	٢
ثالثاً : الإلمام بالإدارة المدرسية : الإلمام بالشئون المالية والإدارية إلى الدرجة التى تعينه على أداء أعماله دون خطأ مالى أو إدارى . وإسهامه فى الأعمال الإدارية المدرسية :	(ا) يلم عن طريق الممارسة والدراسة بأسس الشئون المالية والإدارية المدرسية . ويستخدم هذه الدراية متطوعا لأداء الكثير من الأعمال المدرسية	٨
	(ب) يلم بالشئون المالية والإدارية بدرجة تسمح له بالاسهام فى الأعمال المدرسية	٦
	(ج) يتطوع أحيانا فى أداء الأعمال المدرسية مستعينا فى ذلك بما يعرفه عن الشئون المالية والإدارية	٤
	(د) معرفته بالشئون المالية والإدارية قليلة مما يقلل إسهامه فى الأعمال المدرسية	٢

ويرسل الدليل إلى المدرس الأول للمادة ، مع ناظر المدرسة اللذان يقومان بتقدير المدرس . ولكل فقرة وزن كما هو مبين تحت فئة « الدرجة » وتجمع هذه الأوزان وتوضع ، بالطريقة الآتية :

موضوع التقدير ودرجات المراتب المناسبة للمدرس

التسلسل	الصفات	الدرجة
أولا :	الإلمام بالمادة والاستزادة فيها	
ثانياً :	الإلمام بالعملية التربوية والطريقة	
ثالثاً :	الإلمام بالإدارة المدرسية	
رابعاً :	السلوك الشخصي والسمعة الخلقية	
خامساً :	استعداده الذهني والقدرة على التصرف	
سادساً :	ضبط النفس والاتزان الانفعالي	
سابعاً :	علاقة المدرس بالزملاء	
ثامناً :	علاقة المدرس بالتلاميذ	
تاسعاً :	أثره في نواحي الحياة المدرسية	
عاشراً :	نشاطه الاجتماعي خارج المدرسة	
حادى عشر :	المواظبة	
ثانى عشر :	المظهر الخارجي	
	مجموع الدرجات	

المجموع بالأرقام

الدرجة بالحروف

....

اسم المقدر

....

....

التوقيع

....

....

• الوظيفة أو المرحلة التعلمية التي يرى ترقية المدرس إليها .
• الوظيفة أو المدرسة التي يرى المقدر نقل المدرس إليها .
• ملاحظات المقدر التفصيلية عن سلوك المدرس وسمعته الخلقية .

ويقوم المقدر بكتابة ملاحظاته في مكان مخصص لهذا في الدليل . وكل هذه الخطوات متبعة في كل الأدلة بما في ذلك عدد الفقرات التي يقدر على أساسها المشتغل بالتربية ، وكذلك أوزان الفقرات وطريقة صياغتها وخطوات

الإجراء و . . الخ . فمثلا نجد الفقرات الثلاث الأولى فى دليل تقويم الإخصائى الاجتماعى هى :

موضوعات التقدير	مراتب التقدير	الدرجة
أولا : القدرة على توجيه النشاط الاجتماعى العام فى المدرسة : وضع خطة سنوية للنشاط الاجتماعى العام بالمدرسة والمرونة والمهارة فى تنفيذها ، وتوجيه نواحى النشاط الاجتماعى من حفلات وخدمات عامة بحيث تؤدى إلى الأغراض التى أقيمت من أجلها .	(ا) ذو مقدرة كبيرة على تنظيم الحفلات والندوات وتوجيهها بحيث تؤدى إلى ما يقصد منها	١٢
	(ب) يحكم وضع الخطة السنوية للنشاط الاجتماعى العام بالمدرسة ، وبارع فى تنفيذها حسب ظروف المدرسة ، ويحرص دواما على أن تؤدى أغراضها	٩
	(ج) يعمل على تنظيم وتوجيه نواحى النشاط الاجتماعى العام بالمدرسة	٦
	(د) اهتمامه قليل بوضع خطة واضحة للنشاط الاجتماعى العام بالمدرسة بحيث يرتبك أمام ما يطرأ على المدرسة من ظروف	٣
ثانيا : القدرة على العمل مع الجماعات المدرسية المختلفة . تعاونه مع الجماعات المدرسية فى النواحى الاجتماعية والثقافية والفنية والصحية واشتراكه فى المجالس المختلفة بالمدرسة .	(أ) يتطوع بشكل فعال فى نشاط الجماعات والمجالس المدرسية ، ويظهر أثر ذلك فى نتائج هذا النشاط	٨
	(ب) يبث روح التعاون بين الجماعات والمجالس المدرسية باشتراكه فى أعمال كل منها	٦
	(ج) يتعاون مع الجماعات المدرسية المختلفة فى أعمال المجالس المدرسية	٤
	(د) يؤثر بعض الجماعات المدرسية بنشاطه المحدود وقلما يشترك فى أعمال المجالس المدرسية	٢

موضوعات التقدير	مراتب التقدير	الدرجة
ثالثاً : الإلمام ببيئة المدرسة وخدمتها : الإلمام بالبيئة المحيطة بالمدرسة وإسهامه وتلاميذه فى خدمتها والاستفادة من إمكانياتها فى خدمة المدرسة .	(أ) يتقن معرفة البيئة المحيطة بالمدرسة مما يعينه على فهم حاجاتها وبذل جهده فى خدمتها والاستفادة من إمكانياتها فى خدمة البيئة	٨
	(ب) لديه المعرفة الكافية بمشكلات البيئة ويعمل على حلها ويستعين ببعض إمكانياتها فى خدمة المدرسة	٦
	(ج) يلم ببعض جوانب هذه البيئة ويسهم أحيانا فى خدمتها وقلما يفيد المدرسة من إمكانياتها	٤
	(د) معرفته بهذه البيئة قليلة وخدماته محدودة	٢

ومن المتوقع أن ينتهى إعداد هذه الأدلة بصورتها النهائية ، بعد التجريب ، قريبا .

اختبار الاستعداد للوظائف التعليمية

قامت الدكتورة رمزية الغريب بكلية التربية باقتباس وتمصير هذا الاختبار عن اختبار ستانفورد للاستعدادات التعليمية Stanford Educational Aptitude Test تصميم ميلتون چنسون Milton B. Jensen سنة ١٩٢٨ .

والاختبار يتكون من :

القسم الأول : ترتيب المهنة المفضلة وزمنه المتوسط ٨ دقائق . ويقوم على أساس الاختيار من إجابتين وعدد فقرات هذا القسم ٤ فقرات فى كل يفترض أن أمام المفحوص فرصتان يستطيع اختيار إحداهما . مثل :

المقارنة الأولى : (علّم على واحدة) .

وظيفة رقم ١ : وظيفة فى مدرسة خاصة فى مدينة الاسكندرية بمرتب شهرى قدره ٢٤٠ جنيها مع زيادة سنوية لا تتعدى ٢٤ جنيها . وهناك فصول ليلية بمرتب إضافى لمن يرغب فى ذلك ولا تسمح جداول المعلمين بالدراسات العليا .

وظيفة رقم ٢ : وظيفة فى مدرسة خاصة فى مدينة الاسكندرية بمرتب سنوى قدره ١٨٠ جنيها وتحدد الترقية إلى الوظيفة الإدارية فى الغالب بالأقدمية ، كما أعدت جداول المعلمين بحيث تسمح لهم بالالتحاق بأقسام الدراسات العليا بالجامعة ، ويوجد قسم للأبحاث بمنطقة الاسكندرية التعليمية تشرف عليه جامعة الإسكندرية إشرافاً جزئياً ، ومعظم بحوثه متعلقة بالمناهج ومشاكلها ومستعد لبحث أية مشكلة أخرى متعلقة بالحياة المدرسية إذا توفرت الإمكانيات .

القسم الثانى : مشاكل النظام المدرسى وزمنه المتوسط ١٥ دقيقة .

ويقوم المفحوص باتباع التعليمات : تخير نوعا من أنواع التصرفات المعطاة فيما بعد تطابق ما قد تفعله فى كل موقف وارسم دائرة حول الحرف الذى يدل عليها (ا ، ب ، حـ ، د ، هـ ، و) على أساس الإختيار من عدة إجابات هى :

(ا) أقول له (لهم) مدى خطئه فيما فعل .

(ب) لا أفعل شيئاً .

(حـ) أنزل به عقاباً جسمانياً .

(د) ألومه (ألومهم) أمام جميع التلاميذ .

(هـ) أحرمه (أحرمهم) من بعض الحقوق .

(و) أطلب منه (منهم) أن يعدنى بألا يعود إلى هذا العمل .

وفى التعليمات يطلب تصور المفحوص أنه مدرس ويقرأ . المشاكل وعددها ٢٦ حالة . وبعد أن يختار متغير الإجابة من الإجابات الست السابقة يحدد درجة تأكد منها متبعا التعليمات الآتية :

– بعد اختيارك مباشرة لنوع التصرفات المناسب لك فى كل موقف ، بين مدى ثقتك بهذا الاختيار لكى يتبين إلى أى مدى أنت متأكد من اختيارك . ارسم دائرة حول الرقم المقابل لدرجة اختيارك (١ ، ٢ ، ٣ ، ٤) .

درجات التأكد :

١ – متأكد تماماً .   ٢ – متأكد .

٣ – أكاد أشك .   ٤ – غير متأكد تماماً .

القسم الثالث : نشاط المدرسة الثانوية ومتوسط زمنه ١٠ دقائق ويفترض أن المفحوص يتصور نفسه دائماً ناظر مدرسة ثانوية ، عند الإجابة باختيار واضعاً دائرة حول رمز إحدى خمس متغيرات وبيان التأكد بوضع دائرة حول رقم إحدى أربعة حالات بنفس الطريقة السابقة . وهنا نعرض له ٢٦ وجها للنشاط ونأخذ رأيه فيما يتوقع أن يفعله بالنسبة لكل نشاط منها .

وللاختبار مفتاح يحتاج إلى بعض التدريب ممن يستخدمه وهو يقيس استعداد المفحوص للعمل بالتدريس أو البحث أو الإدارة . ويجرى الآن إعداد الاختبار بالصورة النهائية بما يجعله مناسباً لبيئتنا وحاجاتنا .

الفصل الثاني عشر

تقييم التلميـــذ

- ماهى الدرجة .
- وظائف الدرجات والتقديرات .
- وسائل الإتصال بين المدرسة والبيت .

الفصل الثانى عشر

تقييم التلميـــــذ

ما هى الدرجة :

درجة التلميذ فى المدرسة أو المعهد أو الجامعة ، فى نهاية الأمر ، هى حكم يصدره عليه شخص آخر . وقد يكون الحكم حدسيا أو ذاتيا ، وقد يكون موضوعيا تماما عن طريق إجراء الاختبارات المقننة مثلا . وفى الحالة الثانية نجد أن :

١ ـ الاختبارات الموضوعية تقدم أسسا مناسبة للوصول إلى نظام للترتيب.

٢ ـ أن فقرات معينة قد دلت عن نواحى معينة فى التحصيل :

٣ ـ أن الاستجابات المناسبة عن هذه الفقرات كانت هكذا : : : (كما حدث فعلا) .

٤ ـ أن نسبة مئوية معينة من الجماعة يمكن أن تعتبر فى المستوى ا

وهكذا نرى أن الطرق التى نصل بها إلى التقديرات النهائية مختلفة . فقد يدخل الحكم أساسا فى تخطيط عمليات الاختبارات الموضوعية ، وقد يصدر على الفصل ككل ، وقد يدخل فى تقدير أداء تلاميذ معينين . والصفات التى قدرت والأوزان التى حددت لكل تلميذ قد تختلف اختلافا كبيرا من مدرس إلى آخر ومن فصل إلى آخر . ولكن الدرجة ، على أى حال ، هى حكم يصدره المدرس على التلميذ .

والدرجة حكم نسبى Relative فإذا كانت درجة التلميذ مثلا ٧٥ ٪ فإنها قد تكون درجة عالية جدا ، أو ليست عالية ، والأمر يتوقف على

طبيعة الاختبار والعلاقة بين الاختبار والشخص أو الجماعة المختبرة . كما تختلف دلالة الدرجة من مدرس إلى آخر ومن مادة إلى أخرى .

والاختلاف من مدرس إلى آخر ومن مكان إلى آخر يشير إلى أنه ليس للدرجات نقطة مرجعية . وقد يكون المستوى المرجعى هو الجماعة نفسها كأن يقسم تلاميذ الفصل إلى ٢٠ ٪ ا ، ٣٠ ٪ ب ، ٤٠ ٪ ج ، ١٠ ٪ د مثلا . والمعيار Standard قد يكون انطباع المدرس عن أداء الفصول الأخرى وما يتوقعه منطقيا من إنجاز الجماعة المعينة على أساس أنها وصلت إلى مستوى معين من التقدم . وفى كلتا الحالتين المعيار Standard نسبى . فليس هناك أى مستوى ثابت نرجع إليه أداء الفرد .

ولا تستطيع المدرسة أن تتهرب من إصدار مثل هذه الأحكام النسبية على التلاميذ . فهذه الأحكام ستكون هامة دائما فى إرشاد التلميذ فى عمله المدرسى وفى فهم عاداته الشخصية ونقائصه ، وفى مساعدته فى تخطيط مستقبله التعليمى والمهنى ، وفى التعاون مع المدارس الأعلى وأصحاب العمل على انتقاء الأفراد الأكثر صلاحية . ويجب هنا أن نجيب عن التساؤلات الآتية :

١ ــ متى يجب إصدار الأحكام التقييمية ؟

٢ ــ فى أى صيغة يجب أن تسجل هذه الأحكام ؟

٣ ــ ما هى العوامل التى يجب تغطيتها ؟

٤ ــ على أى أنواع الأدلة يجب أن تعتمد الأحكام ؟

٥ ــ من يجب أن يكون مسئولا عن تقدير الأحكام ؟

٦ ــ لمن يجب أن تبلغ ؟ من هو المستهلك المناسب للمعلومات ؟

وظائف الدرجات والتقديرات

أولا : المعلومات التى يحتاجها التلميذ الذى نقدره

لماذا يحتاج التلميذ إلى أن يعلم تقدير المدرسة له ؟ وما هى المعلومات التى يحتاجها . يبدو أن المعلومات المرسلة للتلميذ تخدم ٤ وظائف : (١) دفع العمل المدرسى (٢) توجيه التعليم (٣) توجيه الخطط التعليمية والمهنية (٤) توجيه النمو الشخصى .

١ – الدفع : الدوافع المتضمنة فى العمل المدرسى كثيرة مركبة وهى تتضمن الرضى عن الإشتراك فى جماعات ، وعن الاستطلاع واستكشاف ميادين جديدة ، والرضى عن النجاح ، والرضى الناتج عن التقبل الاجتماعى ، والرضى عن مسايرة التوقعات و . . الخ .

كما أن هناك حوافز سلبية وهى حقيقية مثل الإحباط والفشل و . . الخ .

الاختبار والتقييم والدرجات ونظام التقارير كلها تدخل إلى دفع التلميذ عن أربع طرق على الأقل هى :

(ا) خبرات النجاح والفشل فى المحفوظات اليومية ، والإختبارات ، والتمارين المدرسية .

(ب) العلم بأن الآخرين يقدرون عمله ويكونون رأيهم عنه على أساس عمله .

(حـ) بأن أدائه أيا كان تقدير المدرسة له سيصبح سجلا دائماً وجزءاً من تاريخه الماضى الرسمى .

(د) توقعه لما ستجىء به التقارير المدرسية وأثرها الفعلى على عالمه ، عالم الأسرة وعالم الزملاء .

ولكن بعض رجال التربية ينتقدون دفع الدرجات المعطاة فى المدرسة بالطريقة التقليدية على أن الدرجات :

(ا) فردية وتنافسية وتؤكد على التحصيل والتفوق الفردى على حساب التعاون والتحصيل المشترك .

(ب) غير جوهرية ولا ترتبط بأى أغراض أصيلة أو بالحاجة إلى تعلم الأشياء لذاتها .

(حـ) وهى حائل دون تنظيم العمل فى المدرسة حول اهتمامات وحاجات التلاميذ الأصيلة .

٢ – توجيه التعلم :

إن المعلومات عن نجاح وأخطاء التلميذ هامة أيضاً فى توجيه نواحى نشاط تعلمه . وقد بين البحث قيمة تعيين أخطاء بذاتها لدى التلميذ ، وأوجه نقصه وقيمة الإجراء الموجه نحو إحلال الأداء الصحيح محل تلك الأخطاء ، والطرق الفنية فى تشخيص نقاط الضعف الفردى لديه بحيث تعينه وهنا تصبح هـذه المعلومات أدوات أساسية فى توجيه عمليات التعلم . والدرجات بالطرق الحالية لا تؤدى وظيفتها فى توجيه التعلم إذ أن درجات الفترات هى ملخص التقديرات بطريقة بعيدة عن التعلم الفعلى لأوجه النشاط وهى ليست كلية ، كما تعوزها القيمة التشخيصية ؛ كما أن الدرجات لا تساعد فى توجيه نواحى النشاط اليومى فى المدرسة .

٣ – توجيه الخطط التعليمية والمهنية التالية :

يحتاج التلميذ كى يضع خططا ناجحة وواقعية عن مستقبله التعليمى المهنى ، بجوار حاجته إلى أشياء أخرى ، إلى فهم واقعى لمستواه فى التحصيل المهنى التعليمى . وهو يحتاج هذه المعلومات حاجة بالغة وقت اتخاذه القرارات ووقت تنفيذها . وعلى أى حال فإن القرارات والخطط تتبلور تدريجياً بحيث أنه فى كل وقت يكون من المناسب أن نعكس للتلميذ الفرد صورة واقعية عن نفسه فى علاقته بتوقعاته التعليمية والمهنية .

ودرجات الفترات ، إذا وضعت بانتظام وبمنهجية تكون أحد طرق إبلاغ الفرد بالأحكام عليه وتصبح ذات دلالة فى تخطيطه التعليمى والمهنى . فإن النجاح الأكاديمى فى لحظــة يبصر الفرد بتخطيطه فى هــذه اللحظة . والمشكلة التى تواجهنا الآن هى ، « ما هى أحسن الطرق » لتقدير التحصيل الأكاديمى وما هى أفضل الطرق لتوصيل هذا التقدير إلى التلميذ .

أما عن طريقة التقدير فيمكن مثلاأن نعتمد على حكم المدرسين أفراداً ، لأنها مبنية على أى أنواع الأدلة يرون أن استخدامها يناسب الوصول إلى أحكامهم هذه . وهناك احتمال آخر هو إصدار حكم ضمنى يقدمه اختبار مقنن معين أو مجموعة من الاختبارات تقدم تقديراً عاماً عن احتمالات النجاح . وللاختبارات ميزة كبرى هى أن مقياس الدرجة له معنى موحد لا يعتمد على مدرس بذاته أو مدرسة بذاتها .

وأما عن إبلاغ التلميذ بتحصيله الأكاديمى عن طريق بطاقات التقارير الدورية فهى طريقة غير ناجحة إذ أنها لا توصل له دلالة هذا التحصيل فى خططه الخاصة . فالأمر يلزمه بعض الترجمة أو التفسير ونحن لا نتوقع أن يقوم التلميذ نفسه بذلك . فالطالب يهتم بدلالة سجله أكثر مما يهتم بالدرجات الخام نفسها سواء أكانت درجات مدرسية أو درجات اختبار مقنن . ويرى بعض المرشدين النفسيين إبلاغ الطالب بالتفسير المشتق من درجاته فقط إذ أن الدرجات نفسها لا دلالة لها عند الطالب . والدرجات فى المدرسة هى مواد خام تحتاج إلى منظم ومفسر ذى خبرة يقوم بدراسة تتبعية غير رسمية لتلاميذه ليعرف النجاح المحتمل فى الجامعة لمن حصل على درجة معينة ، ويقرر الأعمال المعينة التى نجح فيها من حصل على درجة معينة فى المدرسة .

٤ ــ توجيه النمو الشخصى :

نحن نتعلم كيف نعيش مع الآخرين خلال الطريقة التى يستجيب بها الآخرون لأفعالنا . والجماعة التى نعيش فيها هى التى تحدد ما هو السلوك المقبول والمرغوب . وللمدرسة أيضاً مسئولية فى أن تعكس للتلميذ الفرد تقديرها لطرقه فى السلوك ، بحيث يطورها فى اتجاهات مرغوب فيها اجتماعياً وهذا الانعكاس متضمن فى كل اتصال يقوم به موظفو المدرسة مع التلميذ . وكل تعبير عن الاهتمام ، أو الموافقــة ، أو التحمس ، أو الاعتراض هو مرشد يدلل على الطريقة التى يعرف بها مجتمع المدرسة أنماط السلوك المرغوبة وبها يقدر التلاميذ . ونشك فى أن التلميذ يفهم التقديرات الدورية وفى أن لها تأثير كبير عليه . ولكن الاتصال الشخصى المباشر فى المدرسة والتفاعل اليومى وخاصة فى المؤتمرات الخاصة ، هو الطريق المحتمل الذى بواسطته تمارس المدرسة تأثيرها على النمو الشخصى للتلميذ وتطوره .

٥ ــ العلاقة بين تقدير الذات وتقدير المدرسة :

بعض رجال التعليم التقدميين يحبذون أن يكون تقييم التلميذ ، أساساً ، تقييماً ذاتياً . وهم يفضلون أن يضع التلميذ أهدافه الخاصة بنفسه لنفسه وأن يقيم تقدمه نحوها . ونحن نرى أن توجيه نواحى النشاط التعليمى يمكن أن يكون فعالا لو أننا رجعنا إلى الأهداف التى وضعها التلميذ بنفسه .

هذا فضلا عن أنه كثيراً ما يكون تقدير التلميذ لنفسه غير صحيح . والتلميذ العادى تنقصه الأرضية الأساسية للحكم على ما إذا كانت قدراته وإنجازه يحققان له ما خططه من دراسة عليا أو عمل مهنى . ونحن لا نتوقع من الطالب المتقدم لمدرسة أو لعمل أن يحكم على صلاحيته حكماً سليما ، ولذا فإن تقدير الذات ليس بديلا لتقدير المدرسة .

ثانياً : إبلاغ الآباء تقييم الأبناء :

التواصل بين المدرسة والمنزل وبين المنزل والمدرسة هو الأساس فى جعل هاتين الهيئتين تفهمان التلميذ فهما أفضل . وللتقارير التى ترسلها المدرسة إلى الآباء ، ضمن أهدافها ، هدف هو أن يعلم الآباء كيف يبدو التلميذ لموظفى المدرسة فى الموقف المدرسى ، بحيث يستطيع الآباء أن :

١ – يتقبلوا ويؤيدوا ويدعموا الطفل أثناء مواجهته لمشاكل نموه .

٢ – أن يفهموا ويتعاونوا مع برنامج المدرسة من أجل الطفل .

٣ – أن يجعلوا الطفل يتبنى أهدافا مهنية وتعليمية بنائية .

ثالثا : تقديم التقديرات للمؤسسات التعليمية الأعلى :

تحتفظ المدارس بالسجلات من أجل إفادة المدارس الأعلى والمعاهد الفنية التى يحتمل أن يدخلها التلميذ مستقبلا . والمدارس الأعلى والمعاهد عند قبول الطلبة الجدد تطلب هذه المعلومات حتى تضعها فى الاعتبار عند قبول الطلبة الجدد . والمدارس السابقة عليها أن تقدم معلومات صادقة وتنبؤية حتى يمكن أن يقدر النجاح المحتمل للطالب على أساسها .

وما دامت المدارس الأعلى تهتم بالقدرة على التفوق فى الدراسات المجردة ، فيجب أن تهتم تقارير المدارس الأقل مستوى بهذه الناحية . ولكن ما هى أفيد المعاومات وما هى أحسن الطرق لإرسالها ؟

هى المعلومات عن الاستعداد والتحصيل المدرسين. والاختبارات المقننة تفوق تقارير المدرسين المبنية على عينة صغيرة من الأداء فى أن الأولى ترجع إلى جدول عام للمعايير .

أما عن التحصيل فتقديره يحسن أن يجمع بين الناحيتين ؛ تقديرات المدرسين غير الرسمية ، وتقديرات الاختبارات المقننة . فالأولى مرنة ويمكن

أن تكيف بحسب مدى واسع من أوجه النشاط : كالمهارات فى المشاركة الشفوية ، والنشاط فى المعامل ، ومهارات الإنشاء والتعبير ، . . وتقديرات الاختبارات الموضوعية فى هذه المهارات الأساسية هامة ، على الأقل ، كدليل يؤيد تقديرات المدرسين . خاصة وأن الاختبارات المقننة فى مهارات القراءة ، والمهارات الحسابية والرياضية ، ومعرفة العلوم والدراسات الاجتماعية وجد أنها تفيد فى التنبؤ بالنجاح لنجاح فى التعليم الأعلى وأنها تفوق ، فى دقتها ، درجات المدرسين .

وعن التحصيل الأكاديمى فى المؤسسات التعليمية الأعلى نجد أن أغلب المقاييس الناجحة فيه هى التى تعطى تقديراً كلياً يتنبأ بالنجاح الأكاديمى . كما تفيد الاختبارات التى تركز على وظائف معينة إذا كان المطلوب معرفة النجاح المقبل فى ميادين ترتبط بهذه الوظائف . ولكن فى أغلب الحالات نحتاج تقديراً ملخصاً للأداء الكلى ؛ مثل الترتيب المئينى فى الفصل وما إليه من إحصائيات مشابهة .

رابعاً : تقديم التقديرات لأصحاب الأعمال :

كخدمة لكل من التلميذ والمجتمع يجب أن تكون المدارس مستعدة لتقديم المعلومات عن التلاميذ لأصحاب العمل المحليين . والمعلومات التى يحتاجها صاحب العمل تختلف من حالة إلى أخرى . ويقوم ناظر المدرسة أو عضو معين من هيئة التدريس بملء الطلب الذى ترسله مؤسسة التوظيف بدقة وكمال بحسب ما تسمح به المعلومات . وهنا تفيد سجلات المدرسة إن كانت كافية وموضوعية . والمطلوب غالباً هو ملخص تقدير مستوى الأداء .

وصاحب العمل غالباً ما يحتاج معلومات عن نواحى غير أكاديمية للطالب . وربما كانت أهم المعلومات تلك المبنية على حقائق عن النشاط

الخارجي للتلميذ وعضويته فى الأندية والمكاتب التى اشترك فيها إلى آخر الأدلة التى تدل على اشتراكه أو مسئوليته . كما يحتاج صاحب العمل إلى تقدير أعضاء هيئة التدريس للطالب وخاصة إذا كانوا يتذكرونه . ولكن الخطوة الأخيرة عيوب هى التحيز والمبالغة والنسيان .

خامساً : استفادة المدرسة :

تساعد التقديرات المدرسة على تحسين عملية التعلم وتوجيه التلاميذ وتفيد الاختبارات المقننة فى إعطاء تقديرات عن مهارات عامة وعن محتوى ميادين بذاتها . وهذه التقديرات تصبح جزءاً من السجل المجمع . والميزة هنا هى أن المدرس يعرف تلميذه جيداً وهو يحتك به فى الفصل ويعرف مستواه من التمارين والتجارب والملاحظات . . الخ . وسجل المدرسة يفيد المدرسة فى علاقتها بالتلميذ فيما يأتى :

١ ــ أنه مستمر ومجمع .

٢ ــ أنه وصفى وتشخيصى

٣ ــ أنه متفهم عام .

٤ ــ أنه منظم .

الأسس الفنية فى إعطاء الدرجات والتقديرات

أولا : العوامل المؤثرة :

تفيد الدرجات فى خدمة الوظائف الآتية : ــ

١ ــ تساعد فى توجيه الطالب ووالديه فيما يتعلق بخططه التعليمية المقبلة .

٢ ــ تساعد المدرسة فى تقرير مدى استعداد التلميذ للدخول فى برامج انتقائية معينة أو مناهج .

٣ ــ تساعد المستويات التعليمية الأعلى فى تقدير احتمال نجاح الطالب فى برامجها .

٤ — تساعد صاحب العمل المحلى فى تقرير مدى صلاحية الطالب لأعمال معينة تعتمد على مهارات أكاديمية .

والدرجة إذا كانت مقياساً خالصاً للجدارة كانت لها وحدة فى المعنى تجعل من الممكن تفسيرها والإفادة منها . ولكن الدرجة إذا اعتمدت على عوامل غير معروفة فى تقدير الجدارة والجهد والاتجاه كانت غير مجدية . والدرجة ليست أداة للنظام تستخدم فى الإثابة أو العقاب ولكنها سجل للتحصيل . والدرجة الجيدة هى التى تقدر شيئاً واحدا يحسن تقرير الجدارة فيه فى ميدان بذاته من التعليم .

ويعتمد تفسير وتحديد معنى الجدارة فى الميدان على المدرس أو مجموعة المدرسين . وأول خطوة فى تقرير ما يجب أن يمثل فى الدرجة على المقرر هى تحديد الأهداف المتعددة للمقرر وتقرير أهمها هنا . ولكن ما أهمية تقدير المعلومات ؟ وفهم المفاهيم والتعليمات ؟ والقدرة على تطبيق المعلومات فى حل المشاكل ؟ والقدرة على تنظيم الأفكار والتعبير عنها ؟ ومهارات المعمل أو المعالجة اليدوية ؟ وما هو وزن كل منها ؟ أما عن تقدير الجدارة فى مقرر فيجب أن يعتمد على تعيين عوامل معينة تحدد الجدارة فى هذا المقرر بالذات .

ثانيا : أنواع الأدلة ووزن كل منها .

ما دام المعلم قد حدد أهدافه وقرر وزن كل هدف فى تقييمه الكلى للجدارة ، وجب عليه أن يحدد ما هى الأدلة التى يجب أن يجمعها لكى تمثل كل هدف وأن يحدد كيف أن الأنواع المختلفة من الأدلة ستوزن فى تقديره المركب . فيجب أن يجمع الأوزان التى يعطيها عن درجة امتحان الفترة والتمارين والامتحانات الدورية والشهرية والتقارير والمشاركة فى المناقشة داخل الفصل . . الخ .

(ا) امتحان نهاية الفصل الدراسى :

يرى من يعطى الوزن الأكبر للامتحان النهائى الذى يعقد فى نهاية المقرر، أن التقدير الجوهرى لكفاءة الفرد فى ميدان من التعلم يجب أن يعطى لما استوعبه التلميذ فى نهاية هذا الفصل الدراسى . بل وربما كان من الأحسن أن يعطى بعد نهاية انتهاء هذا التعليم بفترة من زمن . وهم لا يعطون اعتبارا لكيف أو متى حصل هذا فالعبرة عندهم بما وصل إليه فى النهاية وما بقى من معلومات حتى نهاية المقرر أو بعده بفترة . وهم لايرون أن للتعليم قيمة ما لم يستمر التلميذ مستوعبا أهدافه حتى نهاية الفصل الدراسى على الأقل .

(ب) الامتحانات العامة :

وإن كنا نوافق على أن تقييم الطالب هو تقدير جدارته فى النهاية إلا أننا لا نوافق على أنه التقدير الوحيد ولا على أنه التقدير الأساسى وهذه هى الاعتراضات :

١ – يستحيل تقرير أنواع معينة من الجدارة فى حدود الامتحان الكتابى . والقدرة على إيجاد وتنظيم المواد المتعلقة بالمشكلة ، والقدرة على عرض مهارات معينة – مثل استخدام الميكروسكوب – والقدرة على المشاركة الفعالة فى المناقشة الجماعية أو مشروع جماعى هى أمثلة للنتائج التى لا يلائمها التقدير عن امتحان كتابى .

٢ – عينة السلوك التى يمكن الحصول عليها فى امتحان هى عينة محددة ، وعلى ذلك يكون التقدير أيضاً محدداً . ولكن إذا تضمن التقدير أدلة من مصادر أخرى كان أكثر ثباتا خاصة إذا كان الدليل الإضافى من مستوى عالى كالموجود فى الاختبار . ونحن هنا نهتم بالكم والكيف معا حتى نصل إلى تقدير ثابت بقدر الإمكان .

٣ – العينة المقاسة وقت الامتحان النهائى قد لا تكون عادلة بالنسبة لأفراد معينين . فقد يكون بعض الممتحنين مريضاً ، أو متعباً ، أو واقعاً

تحت ضغط ظروف خارجية ، أو أن مستواه منخفض لأسباب أخرى وقت الامتحان . وأداء التلميذ فى يوم معين وساعة معينة قد يفشل فى تمثيل مستوى أدائه العادى .

٤ ــ الأداء تحت ضغط الإمتحان قد يفشل فى تمثيل جدارة الفرد تحت ظروف مريحة وسوية . والامتحان موقف ضغط . والضغط يزيد إذا كان الامتحان واحداً هاما ، وكانت نتيجته ذات آثار هامة على مستقبل الفرد . ويختلف الناس فى مدى اضطرابهم تحت مثل هذا الضغط فبالرغم من أن تقدير القدرة على الأداء يؤخذ تحت ضغط وأنه يمكن الاعتماد جزئيا على هذا التقدير ، فإنه يصعب تبرير الاستناد عليه تماما وعلى العكس نجد أن الاعتماد على تقديرات ومصادر مختلفة بالإضافة إلى الامتحان النهائى يخفف من ضغط الامتحان النهائى ورهبته .

٥ ــ قد تكون لامتحان نهاية الفصل البالغ منتهى الدقة آثارا ضارة على التعليم ونشاط التعلم خلال العام الدراسى . والعلاقة بين ما يمكن قياسه فى الامتحان وبين أهــداف التعليم علاقة ناقصة . فإن بعض أهداف التعليم لا يمكن قياسها فى الامتحانات . وعند ما تضع هيئة خارجية الامتحان تكون العلاقة أكثر ضعفاً فإن نواحى الجدارة التى تبدو هامة من الأعمال التى وضعتها هيئة الإمتحان قد لا تتعلق إطلاقا بالأهداف التى تبدو واقعية وجوهرية لدى جماعة التدريس . ويبدو هذا صحيحاً عند ما يقوم ممثلون من مستوى أكاديمى أعلى وأكثر تخصصا ومنتقين بعناية عند ما يقوم هؤلاء بوضع امتحانات نهاية الفصل الدراسى فى المستوى الأدنى من البناء التعليمى . مثل هذا الضبط للامتحانات من أعلى يشيع فى بعض البلدان الشرقية إذ قامت الجامعات ، التى لها أبسط وأصغر احتكاك بالسكان ، بضبط الامتحانات التى تمثل المحصلة النهائية للتخرج من المدرسة الثانوية وجعلتها فى صورة أكاديمية تشبه ما بالجامعة . وطبيعة امتحانات نهاية

الفصل حددت طابع امتحانات نهاية العام فى المدرسة الثانوية ، وهـذه بدورها حددت نمط امتحانات آخر العام فى المدرسة الإعدادية . وعلى هذا فإن نمط التقييم وإلى حد كبير نمط برنامج التعليم لنسبة غالبة من تلاميذ المدارس الإعدادية ، وأغلبهم لن يذهب إلى الجامعة ، ضبطها نظام الامتحان الذى وضع لإنتقاء الأفراد الذين سـيتابعون الدراسة فى مستوى الجامعة .

وأى برنامج للامتحانات فى نهاية الفصل سيكون له تأثيره على التعليم الذى وضعت الامتحانات لتقييمه ، هـذا التأثير قد يكون إيجابياً وبنائياً إذا كانت الامتحانات مبنية على تحليل متبصر ينظر إلى الأمام فى أهداف التعليم أكثر مما تخص مدرسة بعينها سنطبق عليها طريقة التقييم هذه . وعلى أى حال فالأغلب أن المدارس تنظر إلى امتحانات نهاية الفصل التى تضـعها هيئة خارجية على أن لها تأثيرا مقيداً ومحدداً يمنع المدرسة من العمل على تحقيق أكثر الأهداف التعليمية جوهرية .

(ج) الامتحانات الدورية خلال الدراسة :

ونقدم هذه الامتحانات دليلا أفضل على تحقيق هـذه الأهداف أكثر مما تستطيعه الامتحانات النهائية . وللامتحانات الدورية خلال الدراسة ميزة تنحصر فى استخدامها فى الأغراض التعليمية والتقييمية . فإذا أظهر الامتحان الدورى عيوبا كيفية فى تحصيل التلميذ ، كان هذا مرشداً يوجه التلميذ فى إعادة مذاكرة مواد الدراسة بطريقة أفضل . وكل الامتحانات تحدث صدمة على التلميذ ولكن صدمة الامتحانات الدورية أقل حدة من صدمة امتحان نهائى واحد . فالتلميذ فى الامتحان النهائى يتصور أن مصيره معلق على نتيجة الامتحان وحده . بينما فى الامتحانات الدورية يعلم التلميذ أن أمامه فرصا كثيرة لتعويض ما فاته .

وقد وجد أن عقد الامتحان بعد الانتهاء مباشرة من الوحدات التى تدرب عليها التلميذ أفضل من عقد الامتحان بعد انتهاء مجموعة وحدات تدريبية مر فيها التلميذ ببرامج مختلفة . ولكن الامتحانات الدورية تقيس الفهم والتذكر المباشر القريب بينما تقيس الامتحانات النهائية القدر الثابت من الفهم والتذكر للخطوط العريضة فى المنهج . ويعيب الامتحانات النهائية أنها تكيف بحيث تقيم الإلمام بالمفاهيم العامة أو التكامل بين أجزاء المنهج فى إطار يشمل البرنامج الدراسى كله . ولهذه النقائص رؤى أن تكون للامتحانات الدورية نسبة من الدرجة الكلية لتقييم التلميذ ، لا أن تكون هى المحك الوحيد .

(د) المقالات والأبحاث والتقارير :

تقيس هذه التقارير التى يعدها التلميذ بنفسه ، خارج الفصل ، عدداً من الأهداف لا تستطيع أن تغطيه الامتحانات الكتابية . فقد يطلب من التلميذ أن يجمع وينظم أدلة وشواهد على صحة فرض من الفروض : فيمنح فرصة للبحث المتعمق نسبيا فى موضوع معين ويعطى وقتا كافيا بطريقة مستقرة على قدر ما يستطيع . وموضوع البحث أو الجهد الخلاق يتكيف بحسب اهتمام الفرد .

وفى مستويات معينة وميادين معينة يكون تقدير هذه المهارات فى اكتشاف الحقائق وانتقاء الأدلة وتنظيم العرض وجعله كاملا ومنطقيا ، والتعبير عن الأفكار بطريقة واضحة يكون أمراً غاية فى الأهمية : ولهذا فإن تقييم الأنواع المناسبة من التقارير المستقلة يجب أن يعطى وزنا جوهريا فى الدرجة النهائية .

وإحساسنا بأهمية مثل هذه التقديرات يجب ألا يصرفنا عن جعلها ثابتة reliable . والصعوبات هنا هى صعوبات امتحان المقال وتزيد فى أن لكل

شخص موضوع وليس كما هو الحال فى امتحان المقال . وثبات تقدير موضوعات الإنشاء الفردية ، والأوراق ، والتقارير منخفض جداً . وهناك مشكلة أخرى فى معرفة كيف أن العمل الذى يتم خارج الفصل يقيم جدارة التلميذ الفرد ، باعتباره عملا متميزاً عن الأسرة ، والأصدقاء ، والعمليات الأكاديمية .

(هـ) المشاركة فى نواحى النشاط الجماعية :

يجب أن يهتم المدرس فى تقديره المتفهم للتلميذ بطبيعة ومدى اشتراك التلاميذ فى أوجه نشاط الجماعات فى الفصل أو المدرسة وأن يهتم بمدى تعاون التلميذ مع الآخرين ودرجة إيجابية الدور الذى يقوم به التلميذ فى مشروعات الفصل أو مناقشاته ، ومدى حذق تعليقاته ومدى عمق فهمه الظاهر لما يشترك فى مناقشته . وينكر البعض قيمة المشاركة فى أوجه النشاط الجماعية عند تقدير التلميذ مدعين أن طبيعة ومدى مشاركة الفصل أمر يتعلق بالمزاج والاتجاهات أكثر مما يتعلق بالمناقشة والتعاون ، خاصة وأنه ليس من السهل أن نقيم معرفة الفرد وفهمه بالاستماع إلى ما يقوله أثناء المناقشة ؛ ولكن تقدير المناقشة يكون أمراً جوهريا عندما تتضمن أهداف التعليم الخاصة إنماء مهارات العمل فى جماعة .

والنشاط فى العمل :

يجب تقدير دراية التلميذ فى استخدام أدوات المعامل وفى استعمال المواد وفى الاستفادة من الأسس النظرية فى التطبيق العملى خاصة فى الدراسات العلمية والعملية .

وسائل الاتصال بين المدرسة والبيت

١ – بطاقة التقرير التقليدية :

إن نظام الدرجات التقليدى فى المدرسة وبطاقة التقرير هى مقاييس ضعيفة لا توصل إلى المستويات الثلاث السابقة ، فالآباء لا يشجعون على الفهم والشـــعور بالألفة لدرجات مثل ٧٥٪ أو ٨٠ / أو ا أو ب أو حـ فهى درجات باردة وهى تقييم غير شخصى وهى تفشل فى إثارة تعاطف الآباء بل وتحطم الروح المعنوية فى الأسرة ، وإن كنا لا نقبل الصورة الكاريكاتيرية كصورة للحياة ، إلا أن علينا أن نتعرف على لب الحقيقة فيها ونقصد بهذه الصورة بطاقة التقرير .

فالتقرير ، الذى يحوى تقديرات عددية أو رموزا أبجدية ، قد يكون ذا قيمة فى مساعدة الآباء على أن يجعلوا أبناءهم يتبنون أهدافا واقعية . فإذا كانت درجات الأبناء منخفضة ، تشجع الآباء على وضع أهداف محدودة لأبنائهم . ولكن الآباء غالباً لا يفهمون دلالة الطريقة المعينة لوضع الدرجات فى المدرسة المعينة وخاصة فى المستويات الأكاديمية العالية . كما أنه ينبغى أن يرفق بتقدير التحصيل الذى وضعه المدرس تقديراً آخر مشتق من مقاييس الاستعداد المدرسى ومن اختبارات التحصيل المقننة .

ويجب أن نعترف بأن ترجمة الدرجات فى المدرسة ودرجات الاختبار المقنن إلى تنبؤ عن التحصيل المقبل ليس أمرا يسهل القيام به ، حتى على الموظفين المدربين المزودين بكل الأدلة الموجودة . فما بالنا إذن بالوالد غير المدرب الذى لا نعطيه إلا بطاقة التقرير المدرسية ؟ إنه قطعا لن يضع حكما سليما . وإذا كان للمدرسة أن تساعد الوالدين فى وضع خططهم عن الطفل بطريقة واقعية ، وجب عليها أن تزودهما بتفسير لقدراته الأكاديمية والمهنية

مبنى على كل الأدلة الممكنة . وهــذا التفسير لا يمكن إرساله فى بطاقة التقرير .

٢ – تعديل عدد ومعنى الدرجات :

وفيه اختصرت فئات الترتيب إلى متوسط وحسن وضعيف مثلا ولكن هذه الطريقة لم تفعل أكثر من أنها أمدت الآباء بمعلومات أقل عن أبنائهم .

٣ – إرفاق الدرجات برتب على عوامل أخرى :

والعوامل الأخرى هى عوامل غير تحصيلية مثل : يتقبل التغيرات – يتقبل قرار الجماعة – يتقبل المسئولية – المظهر – الانتباه – العناية بالمواد – الحرص – النظافة – إنجاز العمل – السلوك – التعاون – يمكن الاعتماد عليه – المجهود . . . الخ .

ولكن يعيب هذه العوامل غموض دلالتها وذاتية مستوياتها وقلة الفرصة لملاحظتها ، وعدم ثبات الحكم . ونحن نجد أن الاختلاف فى تفسير هذه العوامل لا يؤثر وحسب على المدرسين الذين يضعون الرتب عنها ، بل ويؤثر أيضاً فى الآباء الذين ترسل لهم . كما أن المدرس الذى يقوم بهذه التقديرات لا يستطيع استيعاب رتب كل تلميذ على كل عامل . وقياس العوامل الشخصية بهذه الطريقة الذاتية دائما يحيطه الغموض خاصة وأن السمات لا تعرّف بطريقة صحيحة .

٤ – القوائم التفصيلية :

تحاول بعض المدارس أن تضع فى تقريرها المرسل إلى الآباء مزيدا من المعلومات بأن تحلل العوامل والمهارات إلى عدد أكبر من العوامل

المحددة . ومثل هذا التحليل يعتمد على تحليل المدرسة لأهدافها فى كل مستوى ومحاولتها وضع الأهداف فى عبارات واقعية ونوعية . وهذه الطريقة مصممة بحيث تعطى للآباء معلومات عما تحاول المدرسة تحقيقه وعما وصل إليه التلميذ المعين من أهداف هذه المدرسة .

والتقارير من هذا النوع تفضل التقارير التقليدية إلا أن أمامها صعوبات عدة . وأول هذه الصعوبات هو التعقيد فقد يصل التقرير إلى خمسين حكما منفصلا عن كل تلميذ وبهذه الطريقة يكون التقرير مرهقا للمدرس الذى يقوم به من ناحية ملئه كتابة ومن ناحية الدقة فى تذكر المعلومات الكاملة عن كل تلميذ . كما أنه يقلق الآباء خاصة إذا كانوا غير متعلمين .

ولكن التقرير المبنى على تحليل مفصل للأهداف والذى يغطى كثيراً من النقط الخاصة ، تكون أقسام تقييم كل نقطة فيه عادة كافية وبسيطة وهذا التقرير إذا عنى بالاستعداد والتحصيل والإنجاز معاً نجح فى إمداد الآباء بكثير من البيانات المفهومة خاصة إذا كان أمام كل عبارة متغيرات للتقدير محددة لها معناها عند الآباء .

٥ ـ المقابلة بين الأب والمدرس :

عقد المؤتمرات الدورية بين الوالد والمدرس وكل موظفى المدرسة هى أفضل طريقة للاحتفاظ بتواصل فعال بين المدرسة والبيت وهذه المؤتمرات لها المزايا الآتية :

(١) مرونة التواصل : يمكن أن يعدل المؤتمر بحسب حاجات الحالة المعينة . فيؤكد المدرس على ما يريد مناقشته من أمر هذا الطفل المعين .

(ب) فرص التفسير : يمكن أن يدل التقرير ليس وحسب على حقائق منفصلة بل وعن العلاقات الداخلية بين الحقائق وتقييمها . وفيه تناقش الخطط الطويلة المدى على ضوء الأدلة المتجمعة حتى تاريخه .

(هـ) التواصل متبادل : لا تستطيع المدرسة وحسب أن تبلغ الآباء ولكنها أيضاً تتعلم منهم . فحقائق المنزل تلقى ضوءاً على حقائق المدرسة والعكس بالعكس .

(د) إزالة سوء الفهم : إذا بدى أن الوالد لم يفهم رسالة المدرسة أمكن توضيحها ومناقشته فيها . ولكن هناك عقبات تحول دون الاعتماد على المؤتمرات كطريقة أساسية أو وحيدة فى التواصل بين المدرسة والبيت وهـذه العقبات العملية تنحصر فى :

(ا) الوقت : تستغرق طريقة المؤتمرات الفردية ، إذا اتبعت كما ينبغى ، وقتا طويلا إذ أن على المدرس أن يقابل كل والد تلميذ عنده خلال العام الدراسى . وللطريقة ميزة هى أن المدرس يضع فى ذهنه النقط التى ناقشها كما أنه يسجل ملاحظاته فى المقابلة وهذه تعتبر ضوءا يلقى على السجل الدائم للطفل . فسجل النقاط التى نوقشت ، والمعلومات التى حصل عليها المدرس من الوالد ، والإنطباعات عن الوالد والبيت كل هذا يجب أن يكون مصدراً يمكن الرجوع إليه وهذا يعنى أن الزمن المخصص من وقت المدرس لكل تلميذ وهو أقل من ساعتين فى العام الدراسى بأكمله لا يكفى إطلاقا . ولهذا يجب أن يسمح وقت المدرس بهذه المؤتمرات بعد أن يخطط وقت كل المدرسين .

(ب) المهارة : يتطلب القيام بهذه المؤتمرات الفعالة مع الآباء مهارات خاصة من جانب المدرس . فهذا المؤتمر موقف إرشاد ، وهو يتضمن مهارات فى إقامة علاقة طيبة ، وفى الحساسية لمشاعر ووجهة نظر الأب ، وفى القدرة على تكييف الاتصال حسب شخص الأب : هذه المهارات يمكن أن تحسَّن عن طريق الممارسة والتدريب ولكن برغم هذا ستظل هناك فروقا فردية فى المهارة : ونجاح نظام المؤتمر يعتمد على الأمن الشخصى

والحساسية والفهم من جانب المدرسين الذين يعقدون هذه المؤتمرات مع الآباء .

(حـ) تعاون الآباء : يتحمس بعض الآباء للذهاب إلى المدرسة ومقابلة مدرس الطفل أو المرشد النفسى وبعضهم لايتحمس ، لأسباب مختلفة بالرغم من أن التعاون فى المقابلة الأولى أمر هام . ولكن إذا حدث وتمت المقابلة الأولى وأحيط الوالد علما بمعلومات شعر بقيمتها ولم يُجرح شعوره زاد اهتمامه بتلك المقابلات . وبعض الآباء لا يأتى إلى المدرسة بعد المقابلة الثانية أو الثالثة . وهذا النوع من الآباء غالباً ما تكون المدرسة فى حاجة إليهم . وفى هذه الحالة لا تكون المقابلة أداة ناجحة ويمكن الاستعانة بطرق أخرى . أما فى الحالات التى يكون فيها الآباء متعاونين فإن المقابلة تكون أداة قيمة يعول عليها .

٦ ــ الخطابات غير الرسمية من المدرس إلى الأب :

تستغرق الخطابات وقتا أقصر مما تستغرقه الطرق السابقة . ويمكن بها الاتصال بكل أنواع الآباء . والخطابات لا تخلو من المرونة ولكنها تفتقر إلى أثر الاتصال الشخصى وكذا فإنها لا توفر فرصة لتبادل الآراء أو إزالة سوء الفهم . وكتابة الخطابات الجيدة للآباء فن يحتاج إلى مهارة لا تتوفر فى كثير من المدرسين .

الفصل الثالث عشر

اختبارات المقال والاختبارات الموضوعية

- أهمية الامتحانات
- تخطيط الامتحان .
- أنواع امتحانات المدارس .
- قواعد عامة تراعى قبل إعداد الاختبار .
- رسم الخطة لعمل الاختبار .
- إعداد أسئلة الامتحان .
- إعداد الاختبار الموضوعى .
- المقارنة بين الاختبار الموضوعى وامتحان المقال .
- استخدام اختبار المقال وتصحيحه .
- تحديد أهداف تدريس المادة .
- انتقاد نظام الامتحانات .

الفصل الثالث عشر

اختبارات المقال والاختبارات الموضوعية

أهمية الامتحانات المدرسية :

تقييم تقدم التلميذ جانب هام من عمل المدرس ، فالصورة الصادقة عن مستوى التلميذ وكيفية تقدمه أساس فى التدريس الناجح وتؤدى عمليات تقييم التلميذ إلى خدمة عدد من الوظائف نقتصر على أربع هامة منها وهى :

١ ــ الدافع :

يختلف الدافع إلى حد ما من تلميذ إلى آخر ومن فصل إلى آخر ، وتحدد الامتحانات متى يبدأ التلاميذ المذاكرة ، وماذا يجب أن يذاكروا ، وكيف يذاكرونه . والامتحانات التى صممت بعناية وكفاية تستطيع أن تدفع التلاميذ إلى إنماء عادات المذاكرة الجيدة وتصحيح الأخطاء وأن توجه نشاطهم نحو تحصيل الأهداف المرغوب فيها . فعمليات الاختبار تضبط عملية التعلم إلى حد كبير .

٢ ــ التشخيص والتعليم :

تؤدى الامتحانات إلى تشخيص نواحى الضعف وإلى توفير الدليل عن المعارف والمهارات التى يجب توفرها ، فالأسئلة التى يفشل التلميذ فى الإجابة عنها أو التى يفشل عدد من تلاميذ الفصل فى إجابتها تساعد فى تشخيص النقاط التى تتطلب دراسات تالية .

٣ ــ تحديد أهداف التربية :

إن ما يؤكد عليه المدرس فى تقييمه لتلاميذه يدل على ما يعتبره المدرس هاماً . فالمدرس قد يؤكد فى حديثه مراراً أنه يهتم باستخدام

الأدوات والأجهزة والتطبيقات العملية للموضوعات التى درست ، ثم يضع امتحاناً تعتمد الإجابة فيه أساساً على تذكر مادة كتبت تفصيلا فى محاضراته . وعلى هذا فحتى لو أقسم مراراً لتلاميذه باهتمامه بالتطبيق فإنهم سيعرفون أنه لا يهتم بها فعلا كما ظهر من امتحانه فرجع إدراك التلاميذ هنا هو خبرة الامتحان نفسها . والمدرس قد لا يكون على علم بما يريده لتلاميذه كأن يعتبر الأمر يسير آلياً ولا يناقش نفسه فيما ينبغى أن يحققه من أهداف فى التدريس وبالتالى ما يضمنه من وظائف فى الامتحان .

٤ ـ المقارنة بين التلاميذ والنقل ومنح الشهادات :

المدرس مسؤول عن تقرير مدى استحقاق التلميذ للنقل من فرقة دراسية إلى أخرى أعلى منها ، وعلى هذا فهو يقرر مدى النمو أو التقدم الذى أحرزه التلميذ بعد أن درس برنامجاً معيناً . بل وبيده أيضاً أمر تقرير مدى استحقاق التلميذ لاجتياز مرحلة من مراحل التعليم . وهو هنا أمام أمر أخطر ، إذ أنه عند إعطاء شهادة إتمام مرحلة دراسية يعطى جوازاً بالمرور إلى مرحلة تالية أعلى . هذه المرحلة الأعلى تتطلب بالأغلب لا مستوى أعلى من القدرات المستخدمة فى المرحلة السابقة بل وأيضاً ، وهو المهم ، أنواعاً أخرى من القدرات والمهارات ، ولهذا فعليه أن يكون أكثر حرصاً .

تخطيط الامتحان

١ ـ تحديد الأهداف .

٢ ـ تخطيط المحتوى .

٣ ـ إعداد نسخة من الامتحان .

وتتكون خطة الامتحان عادة من إقرار النقاط التالية :

١ ـ تخطيط المحتوى والأهداف .

٢ ـ اقتراحات خاصـــة عما يمكن أن تغطيه المواءمة بين المحتوى والأهداف .

٣ ـ تحديد النسب المئوية التي سيشغلها كل جزء من محتوى المجال المقاس بالنسبة للامتحان على أساس المحتوى والهدف وتقدير العدد الكلى للأسئلة .

٤ ـ تحديد مدى صعوبة الأسئلة وطريقة تدريجها :

والخطوة التالية هي إعداد أسئلة الامتحان فعلا . وسنناقش هنا اختيار أنواع الأسئلة وتوجيهات لتحسين صياغة استخدام أسئلة المقال . وقبل أن نستطرد في مقارنة اختبارات المقال والاختبارات الموضوعية ننبه القارئ للعود إلى الفصل الثاني .

أنواع امتحانات المدرس

يمكن تقسيم الامتحانات التي يضعها المدرس إلى قســمين كبيرين هما امتحانات المقال أو الامتحانات ذات الإجابة الحرة والامتحانات الموضوعية أو التي تسمى الامتحانات الحديثـــة : ولكل منها مزاياه وقصوره : ولكل موضعه في قياس التحصيل في الفصل .

اختبار المقال :

الخصائص الرئيسية في الإجابة التي يتطلبها امتحان المقال في كل تلميذ هي :

١ ـ أن ينظم إجاباته ببذل أقل جهد .

٢ ـ أن يستخدم ألفاظه الخاصة .

٣ ـ أن يجيب عدداً صغيراً من الأسئلة .

٤ ـ أن يأتي بإجابات كاملة ودقيقة .

الاختبار الموضوعى :

يتضمن الاختبار الموضوعى أو الامتحان الحديث عدداً متنوعاً من طرق الإجابة تشترك جميعاً فى أن الإجابة الصحيحة هى أحد الاحتمالات التى وردت فى إجابات السؤال . وهنا يوضع السؤال ويتبعه الجواب الصحيح ضمن عدة إجابات غير صحيحة أو ناقصة .

وشروط الامتحانات هى :—

١ — أن تكون صادقة أى أن تكون المعلومات التى يختبرها الامتحان هى نفس ما يراد منه اختبارها . فإذا كان الامتحان فى مادة التاريخ وجب ألا تكون للعبارات المنمقــة ، أو جودة الخط ورداءته ، أو الأخطاء الإملائية ، أو الرسم ، . . الخ نصيب فى تقدير الدرجة .

٢ — أن يكون الامتحان ثابت النتائج بمعنى أن لا يتأثر كثيراً بشخص الممتحن أو المصحح أو وقت الإجراء فإذا طبق الامتحان مرة أخرى على نفس الفرد أو طبق امتحان يكافئه من حيث المدى الذى يغطيه ومن حيث صعوبة الأسئلة كانت النتائج متقاربة .

٣ — يجب أن لا يتأثر السؤال بسواه من الأسئلة وأن تكون الأسئلة من الصعوبة بحيث تظهر الفروق الفردية بين الأقوياء ، ومن السهولة بحيث تظهر الفروق الفردية بين الضعفاء . كما يجب أن لا تكون الأسئلة من الصعوبة بحيث ينال فيها تلميذ صفراً ولا من السهولة بحيث ينال البعض فيها الدرجة النهائية . وتقاس صعوبة السؤال بتوزيع درجات التلاميذ عنه ، فإن كانت نسبة من أجابوا عليه إجابة صحيحة نسبة كبيرة كان السؤال سهلا والعكس بالعكس كذلك .

٤ — أن يكون الامتحان شاملا ممثلا للمادة التى يراد تقييم تحصيل التلاميذ فيها .

٥ ــ أن تكون الإجابة موضوعية أى أن تعتمد على الحقائق لا أن تعتمد على ميول ووجدانات واتجاهات التلاميذ .

٦ ــ سهولة الإجراء والتصحيح وقلة التكاليف المالية والإدارية والوقت وإمكان نقل الاختبار من مكان لآخر .

٧ ــ أن يواجه الممتحن نفسه بالسؤال عما يريد تحقيقه من الإجابة عنه بالتحديد وبحيث يتمشى هدف الامتحان مع أهداف التدريس .

٨ ــ ألا تكون الأسئلة ذات أثر صادم كبير وأن يسبق الإمتحان علاقة طيبة لتوفير ظروف الدفع وضمان التعاون ومنع الغش والتزييف .

مزايا الامتحانات الموضوعية :

١ ــ تمنع التقدير الذاتى .

٢ ــ تتفادى غموض الإجابة .

٣ ــ تمنع الإجابات الخارجة عن الموضوع .

٤ ــ تشمل مقداراً كبيراً من المادة المراد الامتحان فيها :

٥ ــ تتمتع بتوزيع سليم لأنواع المعلومات والأسئلة عنها ووزن كل منها فى مجموع الدرجات .

٦ ــ إذا افترضنا وجود خطأ فى بعض هذه الوحدات فإنه يكون خطأ يسيراً جداً .

٧ ــ سهلة للطالب والمدرسة والإدارة المدرسية .

٨ ــ تتفادى تأثير النظافة والاهتمام بالشكليات فى الدرجة .

٩ ــ يسهل بواسطتها عمل معايير .

كيف نضع أسئلة الصواب والخطأ :

١ ــ يجب أن يكون عدد الأسئلة المطلوب الإجابة عنها كافيا .

٢ ـ أن تحتمل العبارة معنى واحداً .

٣ ـ أن لا تكون صياغة السؤال موحية بالجواب .

٤ ـ أن لا يزيد عدد الصواب عن الخطأ زيادة كبيرة .

٥ ـ أن لا يكون ترتيب الأسئلة وتسلسلها مؤديا إلى معرفة الإجابة .

٦ ـ ينبغى تفادى التفصيل فى مواد قاستها مواد أو امتحانات أخرى .

٧ ـ يراعى الاختصار فى عدد الكلمات التى تتركب منها العبارة ما أمكن .

٨ ـ يجب أن تكون العبارة غير معقدة فلا تتضمن أكثر من سؤال واحد أو فكرة واحدة .

٩ ـ يحسن الابتعاد عن العبارات المعممة .

١٠ ـ يعمل تصحيح للدرجة النهائية ضد التخمين وذلك بطرح عدد الإجابات الخاطئة من عدد الإجابات الصحيحة للحصول على الدرجة المصححة .

١١ ـ يحسن استخدام دليل للتصحيح أو مفتاح للتصحيح .

طريقة الاختيار من عدة إجابات :

وتتميز هذه الطريقة بعدة مزايا يدركها من يراجع الفصل الخاص بإنشاء الاختبار ، وهذه المزايا هى :

١ ـ أنها تلفت الطالب إلى ضرورة التميز ومعرفة الحقائق معرفة دقيقة .

٢ ـ أنها تعود الطالب على الحكم الصائب والموازنة وتميز أفضل الأشياء .

٣ ـ من السهل تصحيحها بعد إعداد مفتاح لإجاباتها .

٤ ـ قريبة إلى طابع تفكير الطالب .

٥ ـ يمكن استعمالها فى مواد كثيرة خصوصا المواد الاجتماعية والعلمية .

٦ ـ لا يكاد يدخل فيها عنصر التخمين .

٧ – يستطيع المدرس أن يتعرف منها على الأنواع التى يشيع الخطأ فيها فيستغلها فى تدريسه لمادته وفى وضعه للامتحان التالى .

قواعد عامة تراعى قبل إعداد الاختبار فى صورته النهائية :

١ – يجب إعداد أسئلة أكثر من العدد المحدد للاختبار فى صورته النهائية . كأن نعد ضعف العدد المقدر للاختبار أو ما يقرب من ذلك . لأنه فى كثير من الحالات يتضح ضعف بعض الأسئلة أو عدم صلاحيتها كما يتضح أن بعض الأسئلة لا تتفق مع الأهداف الحقيقية التى أعد لها الاختبار .

٢ – يجب اختبار وفحص الأسئلة جيداً فى فترات متباعدة متعددة فكثيراً ما يبدو السؤال جيداً أو ممتازاً ولكن بعد تركه فترة من الزمن ثم فحصه ثانياً يظهر أن فيه غموضاً أو التباساً أو أنه أصعب مما بدا أولا وهكذا .

كما يجب تجريب الاختبار عدة مرات وتحليل النتائج ومعرفة اتجاهات وآراء المفحوصين وبعض من يلمون بنفس المادة من المدرسين .

٣ – يجب طبع الاختبار بحيث يكون واضحاً (بالرونيو أو البلوظة) وتعد منه نسخ أكثر من المطلوب .

٤ – يجب ترتيب أسئلة الاختبار بحيث يسير فيه المفحوص دون اضطراب أو تعثر ويراعى فى ذلك النقط التالية :

(أ) ترتب الأسئلة بحسب صعوبتها الأسهل أولا ثم الذى يليه فى الصعوبة وهكذا .

(ب) الأسئلة ذات الصيغة الواحدة توضع مع بعضها كأسئلة الصواب والخطأ ثم أسئلة التكميل وهكذا .

(ج) الأسئلة التى تتناول موضوعاً مشتركاً أو تتقارب فى مادتها توضع مع بعضها .

٥ – يجب أن تعد للاختبار تعليمات صريحة واضحة ومحددة وأن تحضر استمارة للإجابة بصورة تبسط عملية الإجابة وأن يعد للاختبار مفتاح .

٦ – يشار فى تعليمات الاختبارات إلى التصحيح ضد الصدفة والتخمين وأن يفهم المفحوص ذلك كى لا يؤخذ على غرة ويكون التصحيح حسب المعادلة الآتية :

$$\text{الدرجة} = \text{عدد الإجابات الصحيحة} - \frac{\text{عدد الإجابات الخاطئة}}{\text{عدد احتمالات إجابة كل سؤال} - 1}$$

رسم الخطة لعمل الاختبار

إن أهم وظيفة لعملية التقييم فى التربية هى أن نعرف إلى أى حد حقق التلاميذ أهداف وأغراض الدراسة . ولكى يقيس الاختبار هذه الناحية يجب أن يوضع له تخطيط دقيق . أما أن يجلس المدرس ويضع عن المقرر عدة أسئلة يضمها فى اختبار فإن ذلك فى الغالب يجعل الأسئلة تقيس المعلومات وتهمل الأهداف . كما أن بعض الموضوعات يسهل كتابة أسئلة عنها بينما يصعب ذلك فى موضوعات أخرى . وفوق هذا كله فإن الأهداف تضيع فى زحمة المعلومات وتواردها وفى اهتمام المفحوص بأيها يختار .

وهذا الكلام ينطبق على اختبارات المقال كما ينطق على الاختبارات الموضوعية . لذلك كان تحديد أهداف تدريس المادة والمدرسة عموماً هو أهم خطوة يجب العناية بها قبل البدء فى عمل الاختبار وكتابة الأسئلة أو اختيارها .

لذلك كان تحديد أهداف تدريس المادة والدراسة عموماً أهم خطوة يجب العناية بها قبل البدء فى عمل الاختبار وكتابة الأسئلة أو اختيارها .

ويجب أن يستعين المدرس على تحديد الأهداف بالبحوث والكتب التربوية عن مادته كما يستعين بمدرسين آخرين وفوق هذا يجب بذل جهد من الدراسة والتفكير لمعرفة هذه الأهداف وتحديدها وترتيبها تبعاً لأهميتها ومعرفة نواحى النشاط والعمليات والجهد الذى تتضمنه هذه الأهداف .

إعداد الأسئلة الموضوعية

فيما يلى بعض القواعد الهامة التى يجب مراعاتها عند إعداد الأسئلة الموضوعية . ولكن تنبغى ألا نعتقد أن مجرد الإلمام بهذه القواعد واتباعها حرفياً كفيل بأن يجعل الأسئلة موضوعية جيدة . فإن إعداد الأسئلة عملية فنية تحتاج إلى ذوق ومران وخبرة بالإضافة إلى مراعاة هذه القواعد : ولكنها تساعدنا على السير فى الاتجاه الصحيح وتجنب الأخطاء .

أولا : سهولة المستوى اللغوى لصيغة الأسئلة : فى جميع الاختبارات التى لا يقصد بها قياس القدرة والفهم والتحصيل اللغوى ينبغى أن تكون الصيغة اللغوية للسؤال سهلة وواضحة بالنسبة للجماعة التى يعد لها الاختبار : والقصد من ذلك هو أن لا تكون قدرة الطالب على القراءة والفهم عائقاً يحول بينه وبين ذكر المعلومات التى يتطلبها السؤال :

أما فى اختبارات القدرة على القراءة والفهم فإن الصيغة اللغوية للسؤال تتدرج فى الصعوبة حتى تصل إلى مستوى أعلى من جميع الطلبة اللذين أعد لهم الاختيار .

ثانياً : لا تضع فى السؤال عبارة بنصها من الكتب المقررة أو نصوص المحاضرات : لأن وضع العبارات الواردة فى الكتب بنصها يشجع الاعتماد على الذاكرة الصماء ويقلل من الاهتمام بالفهم . ويمكن تحقيق ذلك (ا) إما بتغيير مبنى العبارة دون معناها أو (ب) استعمال عبارة أخرى تحمل

نفس المعنى بالضبط . أو (حـ) ذكر إحدى الاستعمالات أو التطبيقات التى تتضمن المعلومات التى يتطلبها السؤال .

ثالثاً : الأسئلة المبنية على الآراء الجدلية ضعيفة : عندما يكون السؤال مبنى على رأى أو قول مأثور عن مصدر خاص يجب ذكر صاحب الرأى أو القول . ذلك لأن الآراء والأقوال القابلة للجدل يختلف الأفراد فى فهمها والحكم عليها . ولكن ذكر صاحب الرأى يكسب السؤال شيئاً من التحديد . على أن الآراء الجدلية غالباً لا تعطى أسئلة جيدة ويجب تجنبها إلا فى بعض الحالات الخاصة التى تقتضى ذلك .

رابعاً : يجب أن لا تكون صيغة أحد الأسئلة تحمل فى طياتها الإجابة على سؤال آخر فى نفس الاختبار . لأن ذلك يجعل الإجابة على السؤال الأخير تعتمد على الحيلة أو نوع من الملاحظة وذلك بالطبع يجعل المقياس غير صالح فى قياس ما أعد له .

خامساً : أن يكون كل سؤال مستقلا عن غيره : أى لا تكون الإجابة على أحد الأسئلة متوقفة على معرفة الجواب الصحيح للسؤال السابق . لأن استقلال الأسئلة يجعل الفرصة متساوية للجميع أمام كل سؤال .

سادساً : أن تكون الإجابات الصحيحة على الأسئلة تسير على نظام عشوائى بحيث لايستطيع الطالب أ نيكتشف أو يعتقد أن الإجابات الصحيحة تسير على ترتيب خاص ثم يحاول أن يبنى إجاباته وفقاً لهذا النظام مثلا :

١ خطأ ٢ صواب ٣ خطأ ٤ صواب ٥ خطأ ٦ صواب

أو ١ صواب ٢ صواب ٣ خطأ ٤ خطأ ٥ صواب ٦ صواب

سابعاً : يجب أن تكون الأسئلة واضحة وصريحة : وخالية من أى خدعة أو فخ . ويجب أن لا يكون السؤال فخاً أو مصيدة هدفها إيقاع الطالب أو خداعه فإن مثل هذا النوع كثيراً ما ينخدع فيه

الطالب الأقوى والأكثر تحصيلا . لأن وفرة معلوماته وثقته بنفسه يجعلانه أقل حذراً فيكون أكثر قابلية للوقوع فى الفخ أما الآخر قليل التحصيل والأضعف ثقة بنفسه يكون متردداً ومتشككاً وبذلك لا يقع فى الفخ وبالطبع تكون النتيجة مخالفة للغرض الأصلى من الاختبار .

لذلك يجب أن تكون صيغة السؤال واضحة وصريحة بعيدة عن العبارات الخداعة والحيل المضللة .

ثامناً : تجنب الغموض والتعقيد : سواء من حيث العبارة اللغوية أو من حيث المعنى الذى يقصد من السؤال .

يجب أن تكون عبارة السؤال واضحة خالية من أى غموض أو تعقيد أو إبهام كما يجب أن لا تحمل العبارة إلا معنى واحداً لا لبس فيه ولا احتمالات .

تاسعاً : تجنب الأسئلة فى نقط تافهة : يجب أن تتناول الأسئلة النقط والموضوعات القيمة وأن تنصب على الأجزاء ذات الأهمية . ويمكن التحقق من ذلك بالتأكد من أن معرفة الإجابة الصحيحة عن هذا السؤال تميز بين الطالب الضعيف والطالب القوى فى المادة التى يقيسها الاختبار .

إعداد أسئلة الاختبار

أسئلة الاختيار من عدة إجابات :

فى هذا النوع من الأسئلة توضع أمام السؤال عدة إجابات ليس بينها إلا إجابة واحدة صحيحة وتكون إجابة المفحوص باختيار أحد هذه الإجابات بوضع علامة أمامه أو تحته . وأسئلة الاختيار من عدة إجابات هى أحسن نوع من الأسئلة الموضوعية ذلك لأنها أكثر مرونة إذ يمكن أن تصاغ بطرق كثيرة ولأنها تصلح لقياس نواحى كثيرة كالتحصيل اللغوى والمفردات ،

والمعلومات العامة والفهم والتطبيقات العملية فهى تصلح لقياس ناحية من نواحى التعلم ما عدا تنسيق وترتيب المعلومات وعرضها وهى أكثر دقة وتقل فيها الصدفة بدرجة كبيرة . ولكن هى فى نفس الوقت تحتاج إلى جهد ومهارة .

ويتكون سؤال الاختيار من إجابات متعددة من جزءين الجزء الأول هو صلب السؤال أو جوهر السؤال وهذا الجزء هو العبارة الأساسية التى تعرض المشكلة المطلوب حلها أو الإجابة عليها وهو قد يكون عبارة ناقصة وقد يكون سؤالا كاملا .

والمبتدئون فى وضع الاختبارات يجدون أن صيغة السؤال الكامل أسهل وهى تجنبهم الكثير من الأخطاء . على أن العبارة الناقصة إذا أعدت بمهارة تكون أقصر وأدل . والجزء الثانى من سؤال « الاختيار من إجابات متعددة » هو قائمة الإجابات التى يختار منها الجواب الصحيح . وكلما زاد عددها فى السؤال كلما ضعف أثر الصدفة والتخمين ويفضل أن تكون هذه القائمة مكونة من أربعة أو خمسة احتمالات . إذ ليس المهم كثرة عدد الإجابات المعروضة فقد يكون عددها خمسة ولكن منها اثنين أو ثلاثة يبدو خطوءها واضحاً فيصير الاختبار الحقيقى من إجابتين أو ثلاثة وتتوقف درجة صعوبة السؤال (الاختيار من إجابات متعددة) على دقة الفروق بين الإجابات المعروضة للاختيار ولكن يكون بينها واحد فقط هو الصحيح .

قواعد تراعى عند عمل أسئلة الاختيار من عدة إجابات :

أولا : يجب أن يكون صلب السؤال واضحاً وأن تبدو جميع الإجابات المعطاة من الجائز أن تكون إجابات على المشكلة الواردة فى صلب السؤال . كذلك يجب أن تكون الإجابات قصيرة وأن لا يكون كل منها شبيهاً بأسئلة الصواب والخطأ .

ثانياً : إجعل صلب السؤال يتضمن أكبر جزء من السؤال لأن هذا يحقق الإيجاز فى الإجابات فيقلل من طول السؤال كله وبالتالى يوفر الوقت والجهد فى القراءة . معنى ذلك أن يكون صلب السؤال طويلا نسبيا بينما تكون الإجابات قصيرة ما أمكن .

ثالثا : تجنب شحن السؤال بأشياء غير ضرورية لأن المطلوب أن تكون المشكلة التى ينطوى عليها السؤال واضحة وبارزة .

إستثناء : يستثنى من هذه القاعدة الأسئلة الموضوعة لقياس القدرة على استخلاص النقط المهمة من وسط مادة كثيرة ؛ فى هذه الحالة توضع فى السؤال مادة غير أساسية مختلطة بالنقط الأساسية أو النقط الهامة .

رابعا : تأكد تماماً أنه لا يوجد بين الإجابات المعطاة إلا جواب واحد فقط هو الصحيح تماماً أو هو الأفضل من بين الإجابات المعطاة .

خامسا : الأسئلة الموضوعة للاختيار والفهم والقدرة على التطبيق (تطبيق المبادئ والمعلومات فى مواقف خاصة) يجب أن تصاغ فى عبارات غير العبارات التى درست بها أو وردت بها فى الكتب والمحاضرات .

سادسا : تجنب الأسئلة التى يكون فيها الجواب الصحيح بينه وبين صلب السؤال ترابط سمعى أو لفظى لأن فى هذه يعتمد المفحوص على هذه الإرتباطات السطحية . وعلى العكس يستفاد كثيراً من هـــذه الترابطات بين صلب السؤال وبعض الإجابات الخاطئة ولكن يجب استعمال ذلك بحذر شديد وعدم الإسراف فيها حتى لا يتحول السؤال إلى فخ مضلل

سابعا : تجنب وجود إشارات نحوية أحيانا : ترد فى صلب السؤال بعض إشارات نحوية يستعين بها المفحوص لمعرفة الجواب الصحيح . كأن ترد فى السؤال صيغة الجمع وتكون الإجابات الموجودة بينها واحد فقط فى صيغة

الجمع . أو كأن يرد فى السؤال ضمير وتكون الإجابات بينها واحد فقط يمكن أن يعود إليه هذا الضمير وهكذا .

ثامنا : تجنب العبارة الطويلة فى صلب السؤال : هناك احتمال كبير أن العبارات الصحيحة تكون أطول وذلك لأن صحة العبارة تقتضى ذكر كثير من الخصائص والتحديدات لقيام صحتها على عكس العبارات الخاطئة فإنها غالباً تكون قصيرة وذلك لأنه لا حاجة إلى الاهتمام بأى تخصيصات أو تحديدات .

أسئلة التكميل والأسئلة ذات الجواب القصير Short Answer & Completion items

الأسئلة التى تكون الإجابة عليها بكتابة إجابة قصيرة (كلمة أو جملة صغيرة) ، تشبه بدرجة كبيرة الأسئلة التى يطلب فيها تكميل العبارة بملء الفراغ المتروك فى العبارة . والفرق الوحيد بين هذه الأنواع من الأسئلة هو أن النوع الأول تكون فيه صيغة السؤال عبارة كاملة ويطلب الإجابة عليها ، أما فى النوع الثانى فتكون العبارة الواردة فى السؤال ناقصة ولا يتم المعنى إلا بوضع الكلمة المكملة لها :

ويصلح هذان النوعان لاختبار تحصيل المفردات واختبار ومعرفة الأسماء والتواريخ والتعاريف وتحديد المفاهيم identification of concepts – وكل المسائل الحسابية والرياضية كالجبر التى تنتهى بجواب عددى أو جبرى محدود .

ولكن هذين النوعين من الأسئلة يتطلبان جهداً كبيراً فى التصحيح أو إعداد مفتاح الإجابة نظراً لتعدد الإجابات وتعدد الكلمات التى ترد فى تكميل الأفراد للجمل وهذا يحتاج إلى دراسة هذه الإجابات المختلفة وتقدير ما يعتبر منها صحيحا .

قواعد تراعى عند عمل أسئلة التكميل والأسئلة ذات الإجابات القصيرة :

أولا : تجنب الأسئلة غير المحددة والأسئلة المفتوحة : الأسئلة المفتوحة والأسئلة غير المحددة تكون لها إجابات كثيرة صحيحة وهذا بالطبع يجعلها سهلة جداً .

ثانيا : تجنب ترك فراغات كثيرة فى العبارة المطلوب تكملتها : يحسن أن يكون الفراغ المتروك مكان كلمة أساسية يعتمد عليها معنى الجملة Key word أما ترك فراغات كثيرة فإنه يجعل الإجابة أسهل كما يجعل الإجابات الصحيحة تتعدد . والأسئلة ذات الفراغات الكثيرة تشجع على التخمين .

ثالثا : إجعل الفراغ المطلوب ملئه قرب نهاية الجملة : عند ما يكون الفراغ المطلوب ملئه قريبا من نهاية الجملة تصبح العبارة ذات معنى محدد نسبياً قبل أن يصل المفحوص إلى الفراغ .

أما جعل الفراغ فى بداية الجملة فإنه يجعل المفحوص يقرأ إلى آخر العبارة بفهم ناقص ويحمله على معاودة القراءة من البداية أو التردد أثناء القراءة وهذه عادات سيئة فى القراءة يحرص المربون على استئصالها .

رابعا : فى أسئلة الحساب والجبر يحسن تحديد نوع الوحدات المطلوبة فى الجواب يحسن تحديد نوع الوحدات المطلوبة فى الجواب على المسائل الرياضية لأن ذلك يجعل عملية التصحيح أسهل كثيراً وأكثر موضوعية كما أنه يبعد عن المفحوص كثيراً من الحيرة والتخمين عن نوع الوحدة التى يكتب بها الجواب . وأحياناً يترك المفحوص كتابة الوحدات أمام الجواب ولا يستطيع المصحح التأكد هل هذا سهوا أم جهلا من المفحوص .

أسئلة الصواب – والخطأ :

كثر إقبال المدرسين على عمل اختبارات الصواب والخطأ ، وذلك يرجع إلى سهولة صياغتها . مما أدى إلى ظهور الأسئلة الموضوعية الضعيفة من هذا النوع . مع أن كثيراً من الأسئلة الجيدة من هذا النوع يعتبر عملية دقيقة ويحتاج إلى دراية وخبرة . ولعل شيوع الأسئلة الضعيفة هذه كان العامل الأول فى كثير من الحملات الموجهة إلى الاختبارات الموضوعية .

والحقيقة أن اختبارات الصواب والخطأ عموما وحتى الجيد منها ، لا تصلح إلا فى العبارات التى لا جدال فى صحتها أو التى يكون خطؤها واضح . ولذلك فهى تستعمل فى الأسئلة التى تتناول معلومات جزئية خاصة أو حقائق قليلة الأهمية . كذلك يمكن استعمالها بنجاح فى اختبار معانى المصطلحات وصيغها ولكن من الصعب جداً استعمال هذا النوع لاختبار الفهم الدقيق والاستنتاج السليم أو لاختبار التطبيقات العلمية .

لذلك كانت اختبارات الصواب والخطأ تتناول جزئيات قليلة الأهمية وتشجع على الحفظ الآلى والدراسة السطحية وتشجع على التخمين وتتأثر كثيراً بعامل الصدفة . كل هذا جعلها هدفا للنقد الشديد . ذلك النقد الذى امتد أثره إلى غيرها من الاختبارات الموضوعية على غير حق .

والصغية الشائعة لاختبارات الصواب والخطأ هى أن تذكر عبارة ويوضع أمامها صواب ، خطأ ويطلب من المفحوص أن يضع علامة خطأ أو علامة صواب على الكلمة التى تمثل حكمه على هذه العبارة . مثال :

فى سنة (١٩٥٨) تم اتحاد سـوريا ومصر لتكوين الجمهورية العربية المتحدة صواب خطأ

وقد أدخلت بعض تعديلات فى صياغة هذا النوع من ذلك وضع خطاً تحت الكلمة أو الجزء الأساسى من السؤال هكذا :

فى سنة (   ) . . . .

ومن تعديلات هذه الطريقة أن يطلب من التلميذ أن يكتب الإجابة الصحيحة عن الأسئلة التى يرى أن إجاباتها المرفقة خاطئة : ويكون التصحيح هنا بكتابة رقم أو اسم أو أى شيء مجرد بدلا من كتابة العبارة الأساسية التى تحتها خط فى السؤال .

وصورة أخرى من أسئلة الصواب والخطأ هى أن ينص على أن العبارة خاطئة فى الجزء أو الكلمة التى تحتها خط ويطلب من المفحوص أن يصحح هذه العبارة بكتابة الكلمة أو الرقم الصحيح فى المكان المخصص لذلك .

قواعد تراعى عند عمل أسئلة الصواب والخطأ :

أولا : تجنب استعمال الكلمات الشاملة لأن المفحوص يستعين بها على معرفة الجواب مثل دائماً ، وأبداً ، فى جميع الأحيان ، لأن المفحوص يرى فى مثل هذه العبارات الشاملة تلميحاً إلى خطأ العبارة وكذلك يرى المفحوص تلميحاً إلى صحة العبارة فى الأوصاف الجزئية مثل فى غالب الأحيان أو أحياناً وذلك التلميح يتجه إليه الطالب الذى اعتاد هذا النوع من الأسئلة وبدأ يكون لنفسه حيلا يستعين بها .

ثانيا : تجنب الصيغ الوصفية غير المحددة : إن الصيغ الوصفية غير المحددة مثل غالبا أو كثيراً أو عظيما أو بدرجة شديدة . يختلف الأفراد فى تفسير مدلولها . فمثلا عندما نقول هذا الأمر يحدث كثيراً فإن تفسير ذلك يختلف فيه الأفراد فمنهم من يعتبره ٩٠ ٪ ومنهم من يعتبره ٨٠ ٪ أو ٧٠ أو ٦٠ أو حتى ٥٥ ٪

ثالثا : تجنب صيغة النفى وبصفة خاصة نفى النفى : إن صيغة النفى كثيراً ما يخطئها القارئ السريع الذى يعنى بقراءة الفكرة لأنه يغفل عن

أداة النفى (لا) ومثيلاتها لصغرها ولأنها ليست من صلب العبارة الأصلية . أى أن المعنى الإيجابى يكون وحدة فكرية كاملة بالرغم من إغفال أداة النفى . أما نفى النفى فهو معقد وكثيراً ما يثير الإلتباس عند المفحوص .

رابعا : تجنب العبارات التى تشمل أكثر من فكرة واحدة : العبارات التى تشمل أكثر من فكرة تقيس القراءة والفهم أكثر من قياسها للمعلومات والتحصيل ولذلك نجدها تصلح جداً لإختبار القراءة والفهم .

على أن هذا العيب يكون أشد إفسادا للسؤال إذا كانت إحدى الفكرتين صحيحة والأخرى خاطئة .

خامسا : تجنب العبارات الطويلة : العبارات الطويلة تحمل فى طياتها تلميحا للإجابة . ذلك لأن العبارات الصحيحة غالبا ما تكون أطول من العبارات الخاطئة . وهذا يرجع إلى ضرورة إشتمال العبارة الصحيحة على بعض التحديدات والقيود كى تصبح صحيحة إطلاقا .

أسئلة التزواج Matching : هى حالة خاصة من أسئلة الإختيار من عدة إجابات والفرق بينهما أنه فى أسئلة الإختيار من إجابات متعددة يكون صلب السؤال عبارة عن مشكلة واحدة ويختار المفحوص الإجابة من قائمة الإجابات المعطاة . أما فى أسئلة التزاوج فهناك قائمتان : القائمة الأولى فيها عدد من المشاكل والقائمة الثانية فيها إجابات هذه المشاكل ولكن بترتيب مخالف . والمطلوب من المفحوص أن يربط كل مشكلة من القائمة الأولى مع جوابها من القائمة الثانية ويكون ذلك بأن يكتب أمام كل مشكلة رقم الإجابة التى وردت فى قائمة الإجابات .

وأسئلة التزاوج تعتبر من الأنواع الممتازة وتفيد كثيراً لاختبار معانى المفردات ، وتواريخ الحوادث ، ونسبة الكتب إلى مؤلفيها ، والأحداث إلى ظروفها أو عواملها ، والنظريات إلى أصحابها والاكتشافات إلى مكتشفيها ،

والرموز مع ما تدل عليه كالرموز الكيميائية وأسماء المركبات التى تدل عليها .

ومن مميزات أسئلة التزاوج أنه يمكن استبدال قائمة الإجابات برسم تخطيطى لخريطة أو رسم بيانى أو رسم لحيوان أو جزء من تشريح حيوان وتوضع أرقام على بعض أجزائه ويكتب المفحوص رقم كل جزء أمام الاسم أو الوصف الخاص به وهكذا .

قواعد تراعى عند عمل أسئلة الربط أو المزاوجة :

أولا : يجب أن تكون جميع المشكلات الواردة بالسؤال متجانسة أى من موضوع واحد أو من طائفة واحدة كأن تكون جميعها مركبات كميائية أو أن تكون جميعها أسماء نظريات علمية أو جميعها معلومات جغرافية وهكذا . وذلك لأن الاختلاف الكبير فى مشكلات السؤال الواحد كثيراً ما تساعد على اكتشاف الإجابة الصحيحة أو على الأقل تخفض كثيراً من صعوبة السؤال وتسمح بالتخمين والحيلة .

ثانياً : يجب أن تكون قائمة الإجابات فى السؤال أكبر من عدد المشاكل الواردة به . لأنه إذا كانا متساويين فإن الإجابة على المشكلة الأخيرة تكون محلولة بنفسها إذ تربط الإجابة الباقية بالمشكلة الباقية .

ولكن يمكن مخالفة هـذه القاعدة عندما تكون الإجابة من قائمة الإجابات ستستعمل أكثر من مرة كجواب لأكثر من مشكلة .

ثالثاً : يجب أن يكون سؤال الربط قصيراً نسبياً وذلك لأن اختبار مشكلات متجانسة يكون ممكنا وأيسر فى القوائم الصغيرة . كما أن القوائم الطويلة تربك وتقع فيها أخطاء كتابية أثناء عملية الربط وهى أخطاء لا علاقة لها بموضوع الاختبار .

رابعا : يجب أن تكون قائمة الإجابات مرتبة ترتيبا منطقيا فمثلا فى قوائم التواريخ أو الأعداد توضع مسلسلة أو فى قوائم الأسماء تكون مرتبة أبجديا

وذلك لتلافى الجهد والأخطاء الكتابية (فى عملية الكتابة) وهى أخطاء لا علاقة لها بالناحية التى يقيسها الاختبار .

خامسا : يجب أن تكون التعليمات واضحة فى شرح أساس عملية الربط كأن يذكر أنه يمكن استعمال الإجابة الواحدة أكثر من مرة .

اختبارات التصنيف : هى صورة من صور اختبار المزاوجة (الربط والتوفيق) matching وهى مفيدة فى اختبار المفاهيم خصوصا ما كان منها متقاربا .

ويتكون اختبار التصنيف من القائمة الرئيسية وهى قائمة تشمل مجموعة من التعاريف والأوصاف أو الخصائص أو التفسيرات وأمام كل واحد منها حرف أو رقم يرمز له . ثم يعطى المفحوص مجموعة أخرى من المفاهيم أو المشاكل ويطلب منه أن يكتب أمام كل منها الرمز الدال على الوصف أو الشرح الذى ينطبق عليها .

قياس الفهم :

الصبغة الغالبة على الاختبارات الموضوعية هى قياس المعلومات ودرجة التحصيل أو حالته : ولذلك كثيراً ما تهاجم بأنها لا تقيس الفهم : والواقع أن الفهم متضمن فيها من ناحيتين : أولا قدرة الفرد على تحصيل الحقائق والمفاهيم تطلب قدرته على فهم ما يقرأ أو ما يسمع . وثانيا أن الإجابة على الأسئلة وعمل المقارنات وإدراك الفروق الدقيقة بين الإجابات المختلفة كلها عمليات تطلب الفهم .

على أن الفهم فى الأسئلة المهتمة بالمعلومات والتحصيل يقاس بدرجة ضعيفة نسبيا . لذلك عملت اختبارات أخرى موضوعية لقياس الفهم .

يتكون السؤال فى اختبار الفهم من فقرة فى أول السؤال يقرأها المفحوص ثم يجيب على أسئلة تتناول المعانى والأفكار والنقد الخاص بتلك الفقرة .

وتقاس درجة المفحوص بعدد الإجابات الصحيحة وهى تتوقف على درجة الفهم كما تتوقف أيضاً على سرعة القراء والفهم فى الاختبارات المعدة لهذا الغرض . ويجب ألا تكون القائمة الرئيسية طويلة وذلك لتحاشى عيوب التذكر والكتابة ويستعمل مثل ذلك أيضاً فى نواحى أخرى غير القراءة اللغوية مثل إعطاء رسم بيانى أو تخطيطى لقياس القدرة على قراءة ما فيه من معلومات وحقائق وكذلك رسم خريطة جغرافية أو رسم منحنى رياضى .

إعداد الاختبار الموضوعى

اختيار نوع وحدات الاختبار : والأنواع هى :

(أ) فقرات تعتمد على الاسترجاع البسيط ؛

(ب) فقرات التكميل ؛

(ح) اختبار الاستجابة من عدة استجابات ؛

(د) فقرات المزاوجة .

اختيار نوع المواد : والمواد غالباً هى :

(أ) لفظية .

(ب) عددية .

(ح) رسوم ؛

(د) رموز ؛

(ه) مكعبات ؛

(و) مواد عملية ؛

(ز) صور أخرى نادرة مثل بقع الحبر .

تأليف الفقرات : وتراعى عادة عنـــد إنشاء فقرات موضوعية الأمور الآتية :

(أ) مراعاة قواعد التعبير اللغوى الجيد :

(ب) استبعاد الكلمات الصعبة .

(حـ) تجنب استخدام ألفاظ وردت بالكتب المدرسية المقررة .

(د) البعد عن الغموض .

(هـ) عدم استخدام فقرات لها إجابات واضحة .

(و) على المؤلف ألا يشير إلى الإجابة أو يوحى بها .

(ز) استبعاد الفقرات التى تعتمد إجابتها على الذكاء وحده .

(ح) استخدام كلمات كمية والبعد عن الكلمات الكيفية .

(ط) البعد عن العبارات المضللة .

(ى) عدم تداخل الفقرات .

(ك) تخصيص مكان للإجابة عن الفقرات فى شكل عمود .

وعند تأليف فقرات الاسترجاع البسيط وفقرات التكميل تراعى الأمور الآتية :

(أ) يكون طول الخطوط المتروكة للإجابة كافيا وموحدا فى كل الفقرات .

(ب) تحديد الاستجابات المرغوب فيها ، كأن تكون كلمة أو تاريخ أو عدد أو رمز أو معادلة أو . . الخ .

(حـ) يجب أن تكون الإجابات المرغوب فيها هامة .

(د) كل إجابة صحيحة تأخذ درجة .

(هـ) غالباً لا يعاقب المفحوص على أخطاء الهجاية .

(و) تنتظم مواضع الإجابة فى نهاية كل جملة .

(ز) تكملة الموضوعات تتبع فى الموضوعات التى تكون كلا موحدا .

(ح) تكملة الموضوعات يجب ألا تعتمد على الإجابة فى مسافات كثيرة

وعند إنشاء فقرات اختبار كاختيار إحـدى استجابتين تراعى الأمور الآتية :

(أ) تجنب استعمال جملتين سلبيتين .

(ب) عدم استخدام جملة تتكون من شقين أحدهما صحيح والآخر خطأ .

(حـ) المحددات التى تعين الاستجابة مثل كلمة دائما وغالباً وأبدا ، تستخدم بحرص شديد .

(د) تطلب الإجابات بطريقة موضوعية جداً .

(هـ) يتساوى عدد الجمل الصحيحة مع عدد الجمل الخطأ .

(و) تخلط العبارات الخاطئة مع العبارات الصحيحة بطريقة عشوائية .

وعند إنشاء فقرات الاختيار من عدة إجابات تراعى الأمور الآتية :

(أ) توضع الكلمات التقديمية فى أول الأسئلة ولا داعى لتكرارها .

(ب) متغيرات الإجابة تصاغ تبعا لقواعد اللغة والأسلوب .

(حـ) عدم وضع اختبارات واضحة الخطأ .

(د) يجب أن تكون لكل فقرة أربع أو خمس متغيرات (إجابات) :

(هـ) توضع المتغيرات دائما فى نهاية كل جملة .

(و) تطلب الإجابات بطريقة موضوعية .

(ز) توزع الإجابات الصحيحة بطريقة متكافئة بالنسبة لترتيب متغيرات الإجابة .

(ح) توزع الإجابات الصحيحه بطريقة عشوائية بين المتغيرات .

وعند إنشاء فقرات المزاوجة تراعى الأمور التالية :

(أ) لكل فقرة إجابة واحدة صحيحة .

(ب) اتساق قواعد اللغة .

(حـ) اتساق التصنيفات .

(د) فقرات المزاوجة لا تكون طويلة جداً ولا قصيرة جداً .

(هـ) توضع الفقرات بطريقة عشوائية فى كل قائمة .

(و) كل مجموعة فقرات مزاوجة توضع فى صفحة واحدة :

(ز) تطلب إجابات موضوعية .

المقارنة بين الاختبارات الموضـوعية واختبارات المقالة :

العامل	المقال	الموضوعية
١ ــ توفر فرصة لاختبار قدرة التلميذ على الانتقاء ، والتنظيم ، أو التكامل	+	
٢ ــ تطلب من التلميذ أن يقـــدم الإجابة لا أن يتعرف عليها وحسب	+	
٣ ــ خالية من عوامل مهارة التعبير وحسن الخط		+
٤ ــ خالية من فرص المراوغة		+
٥ ــ خالية من فرص التخمين	+	
٦ ــ توفر عينة ممثلة كافية من المواد المطلوب قياسها		+
٧ ــ يمكن إعدادها بسرعة	+	
٨ ــ يمكن تصحيحها بسرعة		+
٩ ــ يمكن أن يصحها الكاتب آليا		+
١٠ ــ يكون تصحيحها متسقا بين المصححين المختلفين		+

والموازنة بين أهمية هذه العوامل تختلف من موقف إلى آخر . ومن الواضح أن أيا من النوعين لا يشمل كل المزايا : والمدرس ، فى تقديره لعمل فصله ، يحسن به أن يستخدم طريقتى القياس معاً .

الاستخدام الفعال لامتحان المقال

ستبقى اختبارات المقال مستخدمة فى تقييم أداء التلميذ من أجل المزايا الآتية :

(ا) سهولة إعدادها .

(ب) مزاياها فى تقييم القدرات على تنظيم إجابة عن سؤال .

(ج) وفى تقييم استرجاع recall وانتقاء المعلومات المناسبة .

(د) « « عرض المعلومات عرضاً منطقياً فعالا .

وإذ كان للمدرس أن يستخدمها كان علينا تزويده ببعض المبادئ التوجيهية من مثل ، متى يستخدمها ، وماذا عليه أن يعمل كى يتغلب على نقاط الضعف العام فيها . وهذا الضعف يتضح له جزئياً فى صيغ الأسئلة ، كما يتضح فى عملية تقييم الإجابات التى أجابها التلاميذ .

متى يستخدم اختبار المقال

تعد العوامل التى من أجلها نستخدم اختبار المقال عوامل عملية من ناحية ونظرية أساسية من ناحية أخرى .

اعتبارات عملية مباشرة :

١ ـ يكون اختبار المقال حل وسط عملى فى فصل عدد تلاميذه صغير أو عندما يكون وقت إعداد الاختبار محدداً . فمن السهل إعدادها ، ومشقة الإعداد بالنسبة للاختبار الموضوعى تأخذ وقتاً وجهداً يزيد عما تأخذه قراءة إجابات التلاميذ فى حالة اختبار المقال .

٢ ـ قلة عدد نسخ الاختبار الموضوعى وخاصة إذا كان جهازاً .

٣ ـ عدم وضوح فقرات الموضوعية أحياناً ووجود ما يلتبس على التلميذ فهمه فى أسئلتها . وهذا العيب ليس فى طبيعة الاختبارات الموضوعية ولكنه راجع إلى ضعف إلمام واضعها بفن إعداد هذا النوع . ولكن هذا لا يدفعنا إلى رفضها واللجوء إلى اختبارات المقال بل يدفعنا ، وهذا هو الأفضل ، إلى توضيحها خاصة وأن جهد المدرس هنا سيكون بسيطاً .

إعتبارات نظرية أساسية :

هناك وظائف يجيد اختبار المقال قياسها بأحسن من الاختبار الموضوعى هذه الوظائف هى ؛ القدرة على انتقاء الأفكار ، وربطها ، وتنظيمها ، وخلق أنماط أساسية جديدة من الأفكار ، واستخدام الفرد لغته للتعبير عن أفكاره . ولكن يجدر استخدام الاختبارات الموضوعية فى حالة قياس إعادة صياغة الحقائق العلمية .

يمكن تلخيص عيوب اختبار المقال فى :

١ ـ الفروق فى معرفة الحقائق الأساسية تخفى الفروق فى القدرة على تنظيم هذه الحقائق .

٢ ـ ضغوط الزمن تخفى جودة تعبير الفرد كتابة .

هناك استعمالان آخران لاختبارات المقال ؛ الأول اختبار الكتاب المفتوح open -book T. وفيه يسمح للمفحوص بالإستعانة بكتبه ومذكراته وكل ما يمكن أن يعنيه من جداول . . الخ . وهنا تقل أهمية التذكر ، وتبرز قدرة المفحوص على استخراج المعلومات المطلوبة وقدرته على الانتقاء والاستفادة من الحقائق .

والثانى ، أن يعطى السؤال كامتحان خارج الفصل ودون تحديد وقت

مما يقلل ضغط الشعور بضيق الوقت ويجعله أقرب ما يكون إلى اختبارات القوة من حيث تنظيم المعلومات ومن حيث حسن التعبير كتابة .

تصحيح اختبارات المقال

نذكر فيما يلى بعض خطوات تعين على تقليل أثر الذاتية فى التصحيح وهى تعتمد على تجزئة عملية التقدير إلى نواحى محدده :

١ ــ يقرر المصحح ، مقدماً ، ما هى العوامل التى يبغى قياسها : وإذا كان هناك أكثر من صفة أو خاصية يعمل لكل منها تقديراً مستقلا . فإذا كانت المعلومات مهمة يعمل تقدير منفصل على أساس المعلومات وحدها ، وإذا كان التنظيم مهما يعمل تقدير منفصل على أساس التنظيم : وإذا كانت استعمالات اللغة يعمل تقدير على تقدير الجمل والهجاء : وينبغى أن يحرص المصحح على أن يجعل تقدير أى صفة لا يتأثر بالصفات الأخرى :

٢ ــ أن يعد المصحح ، مقدماً ، نموذجاً للإجابة يبين النقط الهامة ثم يخصص درجة لكل نقطة حسب أهميتها . وبعد ذلك يقارن النموذج بعينة من إجابات التلاميذ ، ثم يعدل النموذج والدرجات المخصصة للنقط الفرعية على ضوء هذه الإجابات . بعد هذا يصبح النموذج معياراً يستعين به عند تقدير كل ورقة .

٣ ــ أن يقرأ المصحح سؤالا واحداً فى جميع الأوراق حتى يتمكن من وضع تقديراته على هذا السؤال على أساس موحد . ثم ينتقل إلى سؤال آخر وهكذا .

٤ ــ يقدر المصحح أوراق الإجابة دون أن يلتفت إلى أسماء أصحابها حتى لا يتأثر بفكرته عنهم ، لتكون تقديراته موضوعية ما أمكن :

٥ ــ يراجع المصحح تقديراته السابقة على أساس الإجابات التى حصلت على تقدير متوسط ، ويعدل فيها على ضوء هذا المتوسط كمعيار .

تحديد أهداف تدريس المادة كخطوة أساسية

لوضع خطة عمل اختبار

نذكر فيما يلى أمثلة لبعض الأهداف الهامة :

١ – يفهم التلميذ معانى المصطلحات والمفاهيم الهامة :
٢ – أن يكون لدى التلميذ محصول قاعدياً من المعلومات والحقائق .
٣ – أن يفهم التلميذ عمليات التطور والعلاقات السببية :
٤ – » » الأسس والقواعد العامة :
٥ – أن يطبق هذه » » » على مواقف جديدة .
٦ – أن يكتسب القدرة على التعرف على النقط غير الأساسية وتفسيرها ؛
٧ – أن يطبق العادات التى اكتسبها فى النقد التقييمى للأدلة .
٨ – أن يظهر مشاعر وتفهمات مرغوبا فيها .

كما أن هناك أهدافاً أخرى من تدريس بعض المواد مثل :

١ – القدرة على استعمال بطاقات الفهارس من المكتبة للوصول إلى مصادر معلومات خاصة :
٢ – استعمال الفهرست لتحديد مواضع المعلومات فى الكتب وفى دائرة المعارف .
٣ – أن يقرأ ويحصل على المعلومات من الرسوم البيانية والجداول والنشرات الدورية والصور :
٤ – أن يكتشف العلاقات ويعمل الاستنتاجات بناء عليها من العرض البيانى أو الأرقام :
٥ – أن يدرك تحديدات مجموعة خاصة من البيانات .

الامتحانات شر لا بد منه

الامتحانات شر لا بد منه للأسباب الآتية :

١ ـ لأنها تختبر وعى المعلومات المطلوبة ؛

٢ ـ لأنها ترفع حماس الطالب ؛

٣ ـ لأنها تدفع التلميذ إلى الاستعداد بالمذاكرة ؛

٤ ـ تحدد الغرض من دراسة المادة .

٥ ـ تقيس كفاية المدرس .

٦ ـ عن طريقها نقيم التلميذ :

٧ ـ منها يمكن تشخيص نواحى الضعف فى برنامج الدراسة ؛

٨ ـ خلالها نستطيع أن نقرر مدى صلاحية المدرس لتدريس مادة بعينها مع تكشف النقاط التى لا يجيد توضيحها .

٩ ـ تعيننا على اختيار أفضل الطريق لتدريس كل مادة :

١٠ ـ تحدد لنا نواحى الضعف والقوة فى المجموعة .

١١ ـ بعد هذا التحديد نستطيع تقسيم المجموعة إلى فصول متقدمة وأخرى متخلفة نتولاها بعناية فى تدريس المادة .

١٢ ـ وتبعا لهذا يمكننا تخطيط برنامج تربوى علاجى يتناول الصعوبات التى يجدها التلاميذ فى المادة تناولا علميا مدروسا .

١٣ ـ نستطيع أن نقرر عن طريقها مدى تحقق برنامج الدراسة لأهداف التعليم .

١٤ ـ لا بد من إجراء الامتحانات لغرض تقييم التلميذ والمدرس والبرنامج والمدرسة وتشخيص نواحى الضعف والقوى فى كل من هؤلاء وتناولها بالعلاج .

١٥ ــ عن طريقها يمكن التنبؤ بمدى النجاح المحتمل للتلميذ المنقول من فرقة إلى فرقة أخرى .

١٦ ــ بواسطتها نستطيع تقرير مدى صلاحية التلاميذ للالتحاق بدراسة أعلى .

١٧ ــ تعين الطالب على اختيار مواد التخصص بعد أن يعرف درجاته فى هذه المواد والمواد القريبة منها :

١٨ ــ تعتبر أداة لتصفية المتقدمين للالتحاق بنوع معين من الدراسات .

انتقاد نظام الامتحانات

١ ــ يوجه أنصار مدرسة الجشطالت إلى الامتحانات الموضوعية نقداً يتلخص فى أن الممتحن لا يهتم بتكوين صورة كاملة عن الطالب ، إذ أنه يقدر الأجزاء كل على حدة . على أن هذا النقد مردود ، إذ أن تكوين فكرة كلية عن الطالب يعتبر انطباعاً inpression ونحن نعلم أن الانطباع يتأثر بذاتية من يكونه وبالمرات الأولى التى التقى فيها مع تلميذه .

إذا أردنا إلغاء الامتحانات التقليدية وجب علينا أن نراعى كثرة الأسئلة ودقتها وكل الشروط التى تتعلق بوضع الأسئلة الموضوعية مع مراعاة تدرجها فى الصعوبة فيأتى الأسهل فى الأول ، على أن نوزع أوزاناً نسبية على هذه الأسئلة بحسب صعوبتها وبحسب أهميتها بالنسبة لأهداف تدريس هذه المادة .

على أن هناك انتقادات أخرى كثيرة توجه إلى نظام الامتحانات نجملها ونرد عليها فيما يلى :

١ ــ أصبحت الامتحانات غاية فى ذاتها بعد أن كانت أداة توصلنا إلى الأهداف . ولهذا يجب ألا ننسى أن الامتحان وســيلة وليس غاية فى ذاتها :

٢ – لم تعد الامتحانات أداة تعليمية والمفروض أنها خبرة يجب أن يفيد منها كل من التلميذ والمدرس والمنهج الدراسى وطريقة التدريس والمدرسة والنظام التعليمى .

٣ – يجب الاعتماد فى كل المستويات عند تقدير الطالب لا على درجة فى الامتحان النهائى وحده بل أيضاً على درجات أخرى أهمها :

(أ) الامتحانات الدورية ويحسن أن تكون شهرية حتى يكون الطالب قد تعلم وحدة كاملة من كل منهج .

(ب) درجات المجهود الشخصى الذى يقوم به التلميذ فى المنزل ؛

(ج) درجات عن اشتراك التلميذ فى المناقشات .

(د) درجات عن النشاط العملى والمعملى للتلميذ ؛

(ه) درجات عن هوايات التلميذ واهتماماته وأوجه نشاطه الأخرى فى المدرسة وخارج المدرسة ؛

(د) تكليف التلميذ بعمل تمارين على كل درس وإعطاءه درجات عنها توضع فى الاعتبار عند تقييمها ؛

(ز) يقوم التلميذ باختيار بحث أو موضوع فى كل مادة يخصص له جزء من الحصص، لمناقشة التلاميذ فى تخطيطه وتوضع عنه درجة لها وزنها فى تقييم التلميذ .

٤ – الامتحانات بصورتها الحالية غالباً ما تؤكد على عامل الحفظ والاسترجاع وتغفل عامل الفهم العام .

٥ – الامتحان النهائى باعتباره الفرصة الوحيدة أمام الطالب وباعتباره أيضاً الحكم الفيصل الذى يقرر أمراً خطيراً فى تاريخ التلميذ المدرسى ، الامتحان بهذا الاعتبار يحدث صدمة تؤثر على استعداد التلميذ له ؛ ولكى نعود التلميذ على مواجهة مثل هذا الموقف يجب إعداده له بأن نعقد

امتحانات تجريبية تماثل موقف الامتحان النهائى من كل الوجوه حتى لا يكون الامتحان النهائى موقفاً غريباً .

٦ ــ قد تؤثر على إجابة التلميذ فى الامتحان النهائى عوامل مختلفة مثل مرضه أو اضطرابه أو حدوث ظروف فى الأسرة مما يجعل اعتبار درجة هذا الامتحان محكاً وحيداً أمراً غير عادل .

٧ ــ أسئلة المقال غالباً ما تشمل عدداً محدوداً جداً من نقاط المنهج ، وهنا يتدخل عامل الصدفة فقد يكون التلميذ من مستوى تحصيلى منخفض عموماً ولكنه أستعد فى نقاط معينة قليلة حدث إن أتت هى أو بعضها فى الامتحان وهكذا لا يكون من العدل أن نحكم بصلاحية مثل هذا التلميذ بينما نحكم بعدم صلاحية زميله الذى ألم بكل المنهج فيما عدا نقط قليلة أتى منها الامتحان ولم يجب عنها بالتالى إجابة كافية . ولعلاج هذا الضعف يجب ، إذا لزم أن نتبع طريقة المقال ، أن نراعى ما يأتى :

(أ) مراعاة أهداف تدريس المادة .

(ب) مراعاة أهداف المدرسة .

(حـ) مراعاة أهداف نظام التعليم .

(د) مراعاة طرق التدريس التى اتبعت .

(هـ) مراعاة النقاط التى أولاها المدرس عناية خاصة والنقاط التى كان لها نصيب أقل من اهتمامه .

(و) مراعاة مستوى التلاميذ .

(ز) تحليل الأهداف التى من أجلها وضع الامتحان تحليلا منطقيا فإن طريقة الامتحان تختلف بحسب أهدافه باعتباره وسيلة تمكننا مرة من نقل التلميذ إلى فرقة تالية ، وفى أخرى تمكننا من الكشف عن الموهوبين . فإذا كان الهدفان طرفى نقيض فيجب بالتالى أن

تصاغ الأسئلة بطريقة تبرز مدى استحقاق التلميذ للنقل إلى فرقة تالية مرة ومدى كفاءته فى الحصول على مكافأة تشجيعية فى أخرى .

(ح) يجب أن نزن كل هذه الأمور فى وضعنا للأسئلة بحيث تتضمن الموضوعات التى يتوقع التلاميذ أن يسألوا فيها بناءا على إعدادهم بطريقة معينة :

(ط) يجب أن تغطى الأسئلة الموضوعات الرئيسية لا التفاصيل الدقيقة التى تتطلب ذاكرة رقية والنقط الهامشية التى قد تخفى على أغلب التلاميذ .

(ى) يجب الإكثار من عدد الأسئلة حتى تكون هناك فرصة للاختيار أمام الطلبة .

(ك) يجب إعداد الطلبة للإجابة عن هذا النوع من الامتحانات كطريقة ترتيب النقط وعرضها والاستعانة بالأمثلة حيث أن كتابة المقال حتى فى العلم هى فن .

(ل) يجب أن لا يتأثر المصحح بلغة الطالب وتنميقه وخطه خاصة فى المواد العلمية والاجتماعية ؛

(م) يدرب المصححون على طريقة موحدة فى تصحيح هذا النوع من الأسئلة :

(ن) تعد نماذج للإجابة عن الأسئلة يراعى فيها مستوى الطلبة والنسبة المطلوب نقلها أو نجاحها فى حدوذ التخطيط العام التعليمى والمهنى :

(س) تصاغ الأسئلة بطريقة لا لبس فيها مع البعد عن الكلمات الغامضة بحيث يكون مضمون السؤال الذى قصده واضعوه مفهوما تماما لدى كل التلاميذ .

(ع) يحسن أن يلحق بالسؤال قائمة بالنقاط التى يجب أن يبرزها التلميذ فى إجابته (مثل مع التأكيد على ... مع الإشارة إلى ... أو موضحا فى إجابتك كذا وكذا .. الخ) .

ولعلاج نقاط الامتحانات والاختبارات التى يضعها المدرس نرى فضلا عما ذكرنا من الأساليب الفنية للتغلب على المشكلات السابقة أن تقوم المدرسة ببرنامج للاختبارات . وسنتناول هـذا بالتفصيل فى الفصل القادم .

الفصل الرابع عشر

برنامج الاختبارات فى المدرسة

- مقـــدمة .
- وظائف البرنامج .
- اشتراك جهاز المدرسة فى تخطيط البرنامج .
- خصائص برنامج الاختبارات الجيدة فى المدرسة .
- برامج مقترحة للاختبارات فى المراحل التعليمية المختلفة .
- السجل المجمع .
- جهود وزارة التربية والتعليم .
- عرض نتائج برامج الاختبارات على الجمهور .
- صعوبات برامج الاختبارات فى المدرسة .
- خطورة برنامج الاختبارات على البرامج التعليمية .
- الاختبارات التحصيلية المستخدمة فى برامج الاختبارات
- الاختبارات النفسية فى البرامج .

الفصل الرابع عشر

برنامج الاختبارات فى المدرسة

مقـــدمة :

الخطوة الأولى فى تخطيط برنامج للإختبارات هى اكتشاف أى أنواع المعلومات عن التلاميذ ترى هيئة التدريس أنها تحتاجها واكتشاف كيف تستخدم نتائج الإختبار . فقبل أن نسأل : « أى الإختبارات يجب أن نعطى ؟ » علينا أن نتساءل : « ما هى المعلومات ، التى ليست لدينا ، والتى نحتاج إليها ؟ » ، « ومتى نحتاج لها ؟ » ، « ومن سيستخدمها ؟ » . فنقطة البدء هى المدرسة ومناهجها ، وهيئة التدريس وحاجاتها . على أن الذى يهمنا ، قبل معرفة أى الإختبارات نطبق ، هو أن نعرف كيف ستستخدم الإختبارات . فالسؤال الأول إذن : لأى الأغراض ستستخدم الإختبارات فى مدرستك ؟ إن تحديد الوظائف والأغراض هو الحصان الذى يجر عربة برنامج الإختبارات .

ملخص وظائف البرنامج

وهكذا نجد أن برنامح الاختبارات فى المدرسة له عدة وظائف يمكن تلخيصها فيما يأتى :

وظائفه فى الفصل	وظائفه التوجيهية	وظائفه التطبيقية
١ ـ تضيف التلاميذ لتلقى التعلم داخل الفصل .	١ ـ إعداد الدليـــل لتوجيه المناقشات مع الآباء عن أبنائهم	١ ـ تكوين وتعيين جماعات الفصول :

وظائفه فى الفصل	وظائفه التوجيهية	وظائفه التطبيقية
٢ ــ توجيه تخطيط نشاط تلامذة معينين .	٢ ــ أن يبنى التلاميذ صورة الذات .	٢ ــ وضع التلاميذ الجدد .
٣ ــ تعين التلاميذ الذين يحتاجون إلى دراسة تشخيصية خاصة وتعليم علاجى .	٣ ــ إمداد التلاميذ بفرص الاختيار الحالية .	٣ ــ المساعدة فى تحديد اللياقة للجماعات الخاصة .
٤ ــ تحديد المستويات التحصيلية المعقولة لكل تلميذ وتقييم التفاوت بين القدرة الكامنة والتحصيل .	٤ ــ مساعدة المدرس المرشد وتطوير فهم الآباء للحالات المشكلة .	٤ ــ المساعدة فى تحديد أى التلاميذ سينقلون إلى فرق أعلى .
٥ ــ تعين الدرجات فى المنهج .		٥ ــ تقييم المنهج الدراسى ونقط التأكيد والتجارب على المنهج الدراسى
		٦ ــ تقييم المدرسين .
		٧ ــ تحسين العلاقات العامة .
		٨ ــ إمداد الهيئات الخارجية بالمعلومات .

أولا : وظائفه فى المدرسة

١ – تقسيم التلاميذ إلى جماعات متجانسة للتعليم داخل الفصل : تفيد الاختبارات المقننة فى معرفة الأفراد ذوى المستوى المتقارب من المهارة الذين يوضعون ، بالتالى ، فى مجموعات تعمل معا على نفس المواد وبنفس السرعة . والاختبارات المقننة تقدم المعلومات بسرعة وبطريقة موضوعية فى بدء العام الدراسى وتمكن من الوصول إلى أقصر الطارق وأسرعها وإلى التعرف على قدرات كل تلميذ ومهاراته . واختبارات القراءة تؤدى هذه الوظيفة فى الفرقة الأولى ، والاختبارات التحصيلية تؤديها فى المستويات الأعلى .

٢ – توجيه تخطيط نشاط تلاميذ معينين : بهذه الطريقة نجعل العمل بموضوعات المهارة يناسب المستوى الراهن لكل تلميذ . فيشجع التلميذ الموهوب على التحرك قدما إلى الأمام بمعدل سرعته هو ، ويزود بأوجه النشاط والإمكانيات التى تعينه على هـذا . ويسمح للتلميذ ذى التحصيل المحدود بأن يتحرك ببطء نحو أهداف أكثر تحديداً . وفى هذا النوع من التخطيط نستعين بمقاييس الاستعداد المدرسى ومقاييس التحصيل التعليمى .

٣ – تعيين التلاميذ الذين يحتاجون إلى دراسة تشخيصية خاصة وإلى تعليم علاجى : عند ما تكون هناك أسبابا تجعل النظام المدرسى محتاجا إلى نوع من الدراسة الشخصية الخاصة لأفراد من التلاميذ وإلى مدرسين معينين كى يوفرا علاجا تعليميا ، هنا يستطيع برنامج الاختبار أن يقدم بيانات هامة تساعد فى تعيين التلاميذ الذين يحتمل أن يفيدوا أكثر من غيرهم من ذلك التعليم . وإذا ترك لمدرس الفصل أن يختار التلاميذ لمثل

هذه الخدمات الخاصة ، تعرض للوقوع فى خطأ اختيار التلاميذ الذين يعتبر مستوى تحصيلهم منخفضا جداً . فإنه يصعب عليه أن يفرق بين القدرة العامة المنخفضة وبين نقص خاص فى مهارة محددة معينة . وبرنامج الإختبار الذى يقدر كلا من التحصيل والاستعداد يساعد فى انتقاء هؤلاء التلاميذ ذوى النقص الذين يمكن علاجهم .

٤ ـ تقييم التفاوت بين القدرة الكامنة والتحصيل : اختيار تلاميذ للعلاج التعليمى هو جزء من مشكلة عامة هى تحديد مدى التفاوت بين التحصيل الممكن والتخصيل الفعلى . هذا التفاوت يساعد المدرس على أن يركز جهوده على تلاميذ معينين فى جماعته الخاصة ، ويساعد فى توجيه المناقشات فى مؤتمر مع الآباء ، أو قد يؤثر على التقدير فى الفترات أو الدرجات الشهرية المرسلة إلى البيوت . والهدف الأخير يهم المدرسة إذ أن بطاقة التقرير تحاول أن تقيم تحصيل التلميذ الفرد فى علاقته باستعداده الكامن . وهنا نحتاج إلى مقاييس جيـــدة لكل من القدرة الكامنة والتحصيل .

٥ ـ تقدير الدرجات فى الفرقة الدراسية : يجب أن يعتمد الكثير من مدرسى المدارس الثانوية على نتائج الاختبارات المقننة كجزء من تقييم أداء الطالب . وقد يلائم هذا مناهج معينة حيث الاختبارات المقننة الموجودة توازى فى محتواها أهداف تعليم المنهج . وعلى أى حال فبرنامج الاختبار على نطاق المدرسة أو المدينة لا يناسب هذا الغرض تماما . فمحتوى الاختبار المقنن هنا هو محتوى عام لا يرتبط بتدريس مدرسة معينة أو فرقة دراسية معينة ، فهو يغطى مدى كليا من المعارف والمهارات . ولهذا السبب ، فالإختبار المقنن لم يعد خصيصا لقياس أشياء معينة تعلمها التلميذ فى فصل معين أو فى فترة تعلمية معينة :

ثانيا : الوظائف التوجيهية لبرنامج الاختبارات

١ – التقرير المقدم للآباء عن تقدم أبنائهم : على المدرسة مسؤولية إحاطة الآباء علما بتقدم أبنائهم . وهذه المعلومات المرسلة يجب أن تتضمن ما إذا كان عمل التلميذ جيداً أم رديئاً . وهذا يساعد المدرس على تزويد تقريره للآباء بالأدلة العملية الموضوعية عن أداء أبنائهم ، مما يبعده عن أن ينظر الآباء إليه نظرتهم إلى المتحيز المتعصب ذى الرأى الفردى الذاتى فقد أصبح مزوداً بالحقائق .

٢ – تكوين صورة – الذات بطريقة واقعية : لبرنامج التوجيه وظيفة عامة هى تحسين فهم كل فرد لقدراته ونواحى نقصه . فبرنامج الاختبار يقدم الدليل الموضوعى عن هذه النقاط من الضعف والتفوق حتى تفسر للتلميذ الفرد .

٣ – مساعدة التلميذ بفرص الاختيار الحالية : أمام التلميذ عادة عدة فرص يجب أن يختار بينها ، فعليه أن يقرر أن يدخل نوعا معينا من التعليم كالثانوى والفنى و . . . وبرنامج الاختبار يوفر دليلا يساعد التلميذ والمرشد النفسى فى إتخاذ هذه القرارات .

٤ – مساعدة التلميذ على تحديد أهدافه التعليمية والمهنية : يساعد برنامج الاختبار التلميذ فى تخطيطه البعيد للتعليم أو العمل . وصورة – الذات التى ساعد فى بنائها تفسير المرشد النفسى لنتائج الإختبار ستؤثر على هذه الخطط .

٥ – تحسين فهم الحالات المشكلة : فى كل مدرسة تلاميذ مشكلين من الناحية التعليمية أو من حيث التوافق الاجتماعى . . . الخ وبرنامج

الاختبارات المنتظم ، بالاضافة إلى النظام الجيد فى التسجيل ، يقدمان بعض الأسس التاريخية لفهم المشكلة الراهنة فلكل مشكلة حالية ماضيها .

كما يساعد البرنامج فى توفير تسهيلات أكثر تعمقا للطفل المشكل . وهنا نجد العديد من الاختبارات التى تناسب حالات فردية وإن وجب أن تتفق مع البرنامج الموحد المطبق على كل التلاميذ .

ثالثا الوظائف التطبيقية لبرنامج الاختبارات

١ ـ تكوين وتعيين جماعات الفصول : إذا كانت المدرسة واسعة بما يكفى لتكوين عدة جماعات فى مستوى الفرقة الدراسية الواحدة أو عدة فصول للمنهج المعين ، وجب أن يتخذ قرار على أسس عميقة يتعلق بمن who يذهب إلى أى which جماعة . والتجميع يصبح فعالا فى حالة المناهج الخاصة المفردة مثل ، الرياضة والإنجليزى ، . . . الخ . ويجب أن يتخذ القرار على أساس الاستعداد أو المستوى الأولى initial للتحصيل .

٢ ـ وضع التلاميذ المحولين من مدارس أخرى : عندما يحول تلميذ من مدرسة إلى أخرى ، يجب أن يتخذ قرار بشأن مستوى الفرقة الدراسية والجماعة التى سيلتحق بها . والاختبار هنا يجب أن يضع فى الاعتبار المستوى التحصيلى فى تحديد الفرقة أو يعدل التعليم فى جماعة الفصل حسب الحاجات المعينة للتلميذ الجديد .

٣ ـ المساعدة فى تحديد اللياقة لمجموعات خاصة : قد تكون هناك فصول للتلاميذ بطيئى ـ التعلم وهنا قد يكون القبول فى هذه الفصول على أساس نقص نسبة الذكاء عن مستوى معين . وقد تكون هناك فصول تدرس مواد كالجبر تتطلب نسبة ذكاء فوق مستوى معين أو تحصيل حسابى معين أو درجة معينة على اختبار تنبؤى معين .

٤ — المساعدة فى تحديد أى التلاميذ يجب أن ينقلوا : التقدير الموضوعى للتحصيل والاستعداد فى اختبارات مقننة يقدم عونا للمدرس كى يحكم على تقدم الطفل مقارنا بتقدم المدرســة حتى يتمكن من نقل الطفل إلى فصل أعلى .

٥ — تقييم المنهج ونواحى التأكيد فيه والتجارب عليه : إن تطبيق نتائج الاختبارات لتقييم منهج دراسى محلى أو التجارب فى منهج دراسى معين يجب أن يكون عادلا . والاختبارات المقننة تقيم فقط جزءاً من من الأهداف التى توصل إليها نتائج تدريس المنهج . وتغطية بقية الأهداف تتطلب طرقا غير رسمية فى التقييم فيجب أن تفسر نتائج الاختبار فى ضوء الصورة الكلية للأهداف الواضحة وغير الواضحة . وعلى أى حال ففى أى برنامج ذى منطق تجريبى يكون من المهم تحديد مدى الكفاية التى استوعبت بها المهارات العامة الأساسية .

٦ — تقييم المدرسين : يضع بعض النظار تقييمهم لجدارة المدرسين أفرادا على أساس نتائج الاختبار المقنن ولكن هذا التقييم وحده لا يعتمد عليه ويعبر التربويون الأمريكيون عن هذا بقولهم **objectives are more than what the test measures.**

٧ — تقييم المدرسة كوحدة : قد يمد برنامج الاختبار ناظر المدرسة بنوع من محك الجودة التعليمية . فإن الجداول الملخصة للمدارس والفصول ببرنامج الاختبارات تنظر لها فى ضوء الاستعداد والخصائص الرئيسية لكل جماعة مما يوضح نقاط القوى والضعف للمدارس والفصول فالمعلومات قد تدل على أن أماكن معينة أو أشخاص معينين يحتاجون إلى مساعدة خاصة أو إلى تغيير فى النقاط التى يركزون عليها . وياحبذا لو كانت الاختبارات

تهدف إلى المساعدة لا الحكم سواء على المدرسين أم المدارس وإلا أصبحت أداة هدم

٨ ــ تحسين العلاقات العامة : قد يرى الآباء أن المدارس تهدف إلى تعليم أبنائهم القراءة والكتابة والحساب ولكن هذه النظرة قاصرة ونتائج الاختبارات الموضوعية تصححها فهي تعرض المهارات والوظائف القريبة والبعيدة التي خدمتها البرامج . فإذا أعلن أن المدارس نجحت في هذه الأهداف أمكن أن تتقدم في نفس الاتجاه .وتزيد أهمية وظيفة الاختبار في العلاقات العامة عندما تلجأ المدارس إلى البعد عن الطريق المألوف المعتاد وإن كانت تصل إلى نفس الهدف بطريقة أثبت العلم أنها أفضل . فهذه الابتكارات تجعل الناس يعترضون أن الأولاد لم يتعلموا المهارات الأساسية . وعلى أي حال فالاختبارات توضح مدى صدق هذه المزاعم ونتائج الدراسات عن تحصيل المدارس في المجتمع المحلى ، أو تقييم تغيرات المنهج ، يجب أن توضع في أيدى المدرسين حتى يمكنهم أن يواجهوا انتقادات المجتمع وأن يقدموا صورة صادقة عن التحصيل في المدارس .

(ا) إمداد الهيئات الخارجية بالمعلومات : هـــذه الهيئات قد تكون اجتماعية أو توظيفية أو مدارس أخرى ومعاهد وبرنامج الاختبار معضداً بنظام كاف في السجلات يقدم المعلومات المطلوبة بطريقة مقننة . هذه المعلومات الموضوعية تتضمن أيضاً الاستعداد والتحصيل .

اشتراك جهاز المدرسة في تخطيط البرنامج

نحن نعتقد أنه في النظام المدرسي يجب أن تكون هناك مشاركة واسعة النطاق في التخطيط الأصلي لبرنامج الاختبارات والمراجعة بين الحين والحين . فمادامت الفئات الأساسية من مستخدمي الاختبار هي : (١) الإدارة والاشراف . (٢) التوجيه ، (٣) مدرسي الفصول . لهذا نعتقد أنه يجب أن يشترك

ممثلون من الفئات الثلاث فى التخطيط فيجب أن نضع فى الاعتبار حاجات كل جماعة منها والزمن اللازم للإجراء والطريقة التى ستعالج بها نتائج الاختبار . ولا يقتصر الاشتراك على تطبيقه وحسب بل والمراجعة الدورية لمعرفة أحسن الطرق لاستخدام نتائج الاختبار :

خصائص برنامج الاختبارات الجيد فى المدرسة

١ ـ اعتبار الوظائف : يجب فى تخطيطنا أن نضع فى الاعتبار الوظائف السابقة والوظائف المحتملة الأخرى فعلى هذا الأساس تختار الاختبارات وتحدد مواعيد تطبيقها بحيث نجمع المعلومات المطلوبة حديثة بقدر الإمكان . فمدرس رياض الأطفال مثلا يحتاج إلى تقسيم الفصول إلى جماعات فرعية على أساس تفاوت سرعات القراءة وهنا يجب أن يطبق اختبار الاستعداد للقراءة فى أكتوبر للسنة الأولى ثم يعاد فى مايو . كما أن التلاميذ والمرشدين النفسيين فى السنة النهائية فى المدارس الثانوية يضعون قراراتهم التى يجب أن تبنى على المعلومات الموفورة عن استعدادات التلاميذ وتحصيلهم ومن المهم هنا أن نحدد ما هى الاختبارات التى سنستخدمها ومتى سنستخدمها .

٢ ـ التكامل : يجب أن ننظر إلى برنامج الاختبار ككل فنربط نتائج الاختبار فى السنة الخامسة بنتائجه فى السادسة بها فى السابعة كذا يجب أن يكون كل نوع من المعلومات المجموعة قد جمع بطريقة وفى وقت يجعله يخدم هذا الكل إلى أقصى حد . ويا حبذا لو استخدمت مجموعة من الاختبارات تصلح لقياس التقدم فى المهارات الأساسية ذلك فى أكبر مدى ممكن وهكذا يكون من الممكن مقارنة الدرجات فى الفرقة المعينة بالدرجات فى الفرقة التالية على أساس أن المحتوى عام . وهكذا تكون الصورة عن نمو التلميذ أكثر صدقا . كما يجب أن نضع فى الاعتبار أثناء قياس المدارس الإعدادية

ما قسناه فعلا فى الابتدائية وما سنقيسه فى الثانوية . فإذا طبقنا اختبار ذكاء فى الفرقة السادسة وجب تطبيق اختبار ذكاء مماثل فى السابعة . وإذا طبقت اختبارات استعداد فى التاسعة وجب تطبيق بطارية مماثلة فى العاشرة .

كما يجب أن يتبع السجل الطالب فى تحركاته من المدرسة الابتدائية إلى الإعدادية فالثانوية ، وتسجل المعلومات فى السجل بطريقة فيها استمرار حتى يكون كاملا بقدر الإمكان . والتكامل أخيرا يعنى تحديد الأوقات التى ستجرى فيهـا الاختبارات بحيث يغطى كل الأغراض ، فاختبار الاستعداد المطبق فى السنة العاشرة مثلا يفيد فى التنبؤ بالنجاح فى المعهد .

٣ ـ الإستمرار : الاستمرار فى البرنامج له ميزتان . فمن ناحية تتجمع البيانات فى السجلات عن التلميذ الفرد فالبيانات عن نتائج الاختبارات التى طبقت على نفس التلميذ منذ سنة أو أكثر تلقى ضوءاً على مركزه الراهن فهكذا نرى ما إذا كان الولد المشكل قد عرض من قبل لاتجاه مماثل أو أن المشكلة ليست جذرية والفائدة الأخرى هى أن استمرار استخدام اختبار أو سلسلة من الاختبارات محليا يسمح بإنماء مستويات توقع محلية ، وقد يأخذ هذا اتجاها غير رسمى للمعايير القومية فى تفسير الأداء المحلى كما قد يأخذ صورة خاصة من المعايير المحلية وعلى ذلك فإذا كان أحد أحياء المجتمع المحلى يتمتع بنسبة مئوية عالية من تلاميذ المعاهد فقد نجد أن المئينات المبنية على المجموع السكانى بمدارس الحى تقدم إطارا أكثر ملاءمة للحكم على المكانة الأكاديمية لتلميذ من تلاميذ الحى أكثر مما يلائمه العمر القومى أو معايير الفرقة .

ومعرفة الاختبارات تتضمن معرفة ما تقيسه كما تتضمن معرفة المستويات المحلية للتوقع . فإذا استخدم المدرسون اختبارات وأفادوا من نتائجها فإنهم سيعرفون ماذا تغطيه هذه الاختبارات وما هى المفاتيح التى

يمكن أن تشتق منها لتشخيص نواحي القوى والضعف فى الجماعة . كما يعرفون ما هى أوجه نقص هذه الاختبارات . ولهذا يحسن أن يستخدم النظام المدرسى فى استمرار نفس هـــذه الاختبارات لفترة من السنين ويتحول عنها فقط عند ما تصبح غير صالحة أو عند ما تثبت الدراسات أن هناك اختبارات أفضل .

برامج مقترحة للاختبارات فى المدرسة

١ ــ برنامج للمدرسة الابتدائية والإعدادية : يتركز اهتمامنا فى هاتين المدرستين على مساعدة الفرد على التمكن من وسائل التعلم والاتصال أثناء تعلمه كيف يعيش ويعمل فى جماعة من رفاقه . وفى نفس الوقت تتكون للتلميذ فردية أساسية من القدرة والمعرفة والفهم فى مستواه الخاص . فالاختبارات يجب أن تغطى هذه الوظيفة فى هذا المستوى .

ويرى البعض أن من اللازم اختبار الأطفال فى الإبصار والسمع ونحن نعتقد أن هذه المقاييس جزءاً من الفحص البدنى وليست جزءاً من برنامج الاختبار المدرسى . ولكننا لا نشـــك فى قيمة الإبصار والسمع وأثرهما فى التحصيل خاصة فى حالة من لديهم عيب فى أحد الحاستين .

(ا) اختبارات القراءة : تعتبر القراءة الطريق الأساسى لكسب كل أنواع المعرفة المنظمة وإن قلت قيمتها فى عالمنا اليوم بسبب انتشار السينما والإذاعة والإذاعة المرئية فالتعلم من الكتب بالرغم من كل هذا سيصبح لب التربية خاصة فى المستويات العليا ولهذا يجب أن نشخص منذ البداية العجز القرائى . هذا ويجب أن يطبق اختبار القراءة كل عام أو عامين ابتداءاً من الســـنة الأولى فى المدرسة الابتدائية حتى نهاية المرحلة الإعدادية .

(ب) اختبار الذكاء الجمعى : للمساعدة فى تفسير نتائج اختبار القراءة والجوانب الأخرى من التحصيل الأكاديمى ، وللمساعدة فى وضع مستويات للتلاميذ ، وللمساعدة فى فهم الحالات المشكلة ، تستخدم نتائج قياس الذكاء ؛ ولاعتبارى الوقت والتكاليف يفضل الاختبار الجمعى وإذا كان سيطبق على السنوات الأولى الابتدائية يجب ألا يكون من نوع الذكاء القرائى ولكن الاختبارات فى هذا المستوى نادراً ما تعطى نتائج ثابتة ؛

(ج) بطارية المهارات الأساسية : وهى تتضمن القراءة وتحل محل اختبار القراءة المستقل ؛ وهى تفيد فى تخطيط البرامج الفردية فى التعليم وفى توزيع التلاميذ على البرامج الخاصة فى التعليم العلاجى .

(د) اختبار استعداد القراءة : قد ننظر إلى هذا الاختبار على أنه دليل لمدرس الفرقة الأولى فى تنظيمه لفصله إلى جماعات فرعية لتعليمهم القراءة وباعتباره أساساً يساعد المدرس على تقييم تقدم التلاميذ أفراداً .

الأنواع الأربعة السابقة من الاختبارات هى أكثرها شيوعا فى هذين المستويين ولكن هناك اختبارات أخــرى تستخدم أحياناً . وترجع قلة استخدامها دائما إلى التكاليف (الوقت ، والجهد ، والمال) فضلا عن أنها تخدم وظائف تبدو أقل أهمية ؛ ومن هذه الاختبارات نذكر :

(أ) اختبارات الذكاء الفردية : إذا كانت الإمكانيات كافية والفريق مدربا كانت أصلح من الجمعية خاصة فى السنين الأولى .

(ب) اختبارات التحصــيل فى محتوى المواد الدراسية : بعض الاختبارات التحصيلية فى السنوات الإعدادية النهائية تعالج محتوى مجالات العلوم والأدب والدراسات الاجتماعية . وبعض الاختبارات التحصيلية تنطبق على مدارس معينة ولا تصلح كجزء من برنامج الاختبار فى المدارس الإعدادية ؛

(ج) أنواع أخرى من المقاييس : وهى قد تستخدم فى حالة تلاميذ معينين كالاختبار التشخيصى إذا كان الطفل متأخرا فى مادة معينة والاختبارات الإكلينيكية كالإسقاطية التى يستخدمها الأخصائى النفسى الإكلينيكى فى دراسة الحالات المشكلة .

ونحن لم نوص باستخدام استبيانات الشخصية من نوع الورقة والقلم بسبب عيوبها فى : (١) صدق المعلومات (٢) عمق التفسير . فإننا نرى أن فهم الطفل كشخص يجب أن يعتمد على الملاحظات المباشرة وغير الرسمية التى يقوم بها المدرس وبقية هيئة المدرسة ؛

٢ – برنامج الاختبار فى المدرسة الثانوية : فى المدرسة الثانوية يصل التلميذ إلى نقطة يواجه فيها عدة فرص تعليمية وقرارات يختار بينها فقد يختار مواد معينة أو مناهج مختلفة ويبدأ فى إعداد خطط المستقبل : هل سيبقى فى المدرسة ؟ أم سيذهب إلى معهد أو ما هى خططه فى عالم العمل خاصة وأن المنهج الدراسى أصبح فى هذه المرحلة يغطى المهارات الأساسية فضلا عن العديد من المحتويات الخاصة ونحن نعطى الأولوية فى هذا المستوى للاختبارات الآتية :

(ا) اختبار الاستعداد المدرسى : يجب أن يتخذ كل فرد قرارا خطيرا يتعلق بالمسافة التى سيقطعها من السلم التعليمى وما إذا كان سيتحول إلى التعليم التجارى أم الزراعى أم الصناعى أم أنه سيواصل الدراسة فى معاهد . واختبار الاستعداد المدرسى يساعد فى الوصول إلى هذه القرارات وله قيمته فى تفسير مدى تقدم التلميذ مدرسيا وفى موضع مستوى توقع التلميذ .

(ب) اختبار قراءة : لا فائدة من اختبار القراءة فى هذا المستوى ما لم تستفد المدرسة من نتائجه فى توجيه النشاط العلاجى ؛ وإذا كان الأمر كذلك وجب تطبيقه فى أول هذه المرحلة ؛

(ه) اختبارات الاستعدادات الخاصة : إذ كنا ننظر إلى التوجيه المهنى خاصة لمن لن يذهبوا إلى المعاهد العليا فإن برنامج الاختبار فى المدرسة الثانوية يجب أن يتضمن مقاييساً للاستعدادات الهامة فى النجاح فى المهن مثل بطارية D.A.T. . واختبارات الاستعدادات الخاصة لا تلائم السنوات قبل الثامنة أو التاسعة ويحدد وقت إعطائها بالبرنامج الكلى للتوجيه فى المدرسة ويجب أن يطبق الاختبار فى الفرقة التى سيتخذ فيها الطالب قرارات هامة .

(د) الميول Interests : الاستعدادات والميول أرضية أساسية للتخطيط المهنى ويمكن قياس الاثنين معاً باختبار مثل اختبار كودر .

(ه) الاختبارات التحصيلية فى محتوى المجال : وهى أقل من سابقيها فى مجال التوجيه وإن كانت ذات قيمة فى تقدير إمكانية مواصلة التعليم .

(و) الاختبارات التنبوئية : تفيد فى مواد معينة مثل الجبر واللغات الأجنبية والاختزال . وتستخدم فى تصفية التلاميذ قبل أخذ هذه المواد فعلا .

(ز) قوائم الشخصية والتوافق : تفيد فى أن يكون فريق الإرشاد أرضية أساسية جيدة وتكون هذه القوائم فى أيديهم أدوات تصفية .

٣ – برنامج الاختبار فى المعاهد : تصبح البرامج هنا متنوعة ومتخصصة بعد أن تكونت المهارات الأساسية . وليس المهم هنا هو التعليم بقدر التوافق والاختيار المهنى كمشاكل أساسية لدى التلاميذ .

(أ) الاستعداد المدرسى : مفيدة جداً خاصة بالنسبة للطلبة الذين يعانون من تحصيل غير مرض لإرشادهم على أساسها . فالتقدير المنفصل للقدرات اللفظية والعددية له قيمة كتقدير تشخيصى . ويجب أن يعطى هذا الاختبار عند بدء المرحلة . ويحسن استخدام بطارية لا اختبار .

ولسنا هنا بحاجة إلى اختبارات الاستعدادات الخاصة لأن المجال المهنى أصبح أضيق من ذى قبل بعد أن قرر التلميذ دخول معهد معين فهو ليس إذن بحاجة إلى اختبارات ميكانيكية ومكانية وكتابية .

(ب) القراءة : هنا نقيس بهدف فهم مشاكل الفشل الأكاديمى : ففى المعاهد أحياناً عيادات قراءة حيث القراءة علاجية . ولهذا فالاختبار يجب أن يطبق فى أول المرحلة لتحديد من سيلقون عوناً خاصاً فى تعليمهم .

(حـ) الميول : هنا ثبتت الميول إلى حد ما واقترب الفرد كثيراً من الاختبارات المهنية النهائية ولذا يحسن تطبيق اختبار الميل فى مهن معينة مثل سترونج .

(د) اختبارات الوضع فى مواد خاصة : تقوم معاهد معينة بعمل دراسات إعدادية على أساس قدرة التلاميذ الجدد . ومثل هذا التقسيم فى الإعداد يوصل المحتويات والتعليم إلى الجماعات المتطرفة . وهنا تطبق اختبارات تقيس مجالا واحداً .

(هـ) قائمة التوافق : تفيد هنا أكثر مما فى المستويين السابقين الأقل نضجاً وتعليما خاصة إذا كانت هناك خدمات إرشادية جيدة . ولكن هناك اعتراض على القوائم والطرق الإسقاطية إذا طبقت جميعا . فاختبارات الشخصية تفيد أكثر لو أتى العميل متطوعاً ، كما أنها تطلب تدريباً خاصاً فى من يستخدمها . .

السجل المجمع

يجب أن تكون المعلومات المتجمعة عن كل طالب فى نهاية الأمر بصورة منظمة بحيث تناسب السجل المجمع Cumultative Record وهو يتضمن كل أنواع المعلومات عن التلميذ وهو أهم ما يفيد المدرسة فى نشاطها التوجيهى للتلميذ وفيه مسئوليتها عن تزويد المعاهد أو أصحاب العمل

بالمعلومات متى يطلبونها . والأغلب أن كل مدرسة تحتفظ ببعض السجلات التى تجمع المعلومات عن الطالب طوال سنى بقائه بها ولكن النظام الحالى لهذه السجلات ليس فعالا بقدر ما نود فهو لا يسمح بإمداد المعلومات النافعة فى توجيه التلميذ فى كل مراحل عمله المدرسى .

وسجلات التجميع شأنها شأن برنامج الاختبار هى أمر محلى يعتمد على الحاجات والإمكانيات المحلية ، فأنواع السجلات التى تحفظ وطريقة ومكان حفظها ، والأشخاص الذين سيحفظونها كل هذه أمور تتقرر بناءا على أغراض وتوجيه المدرسة المحلية . وعلى أنواع المعلومات التى جمعت عن التلاميذ وعلى طرق استخدام هذه المعلومات على إمكانيات التوجيه والمستخدمين الكتبة . . وليس المهم فى هذه السجلات ما نضعه فيها ولكن ما يستنتج منها . كما يهمنا معرفة من سيستخدمها وما هى المعلومات المطلوبة وفى أى الأغراض تستخدم كما يهمنا أن نعرف الأسس العميقة عن محتواها وتكوينها ووضعها وخدماتها .

اعتبارات فى تخطيط البطاقة المجمعة : (أ) ما هى المعلومات التى تتضمنها (ب) فى أى صورة تحفظ السجلات (ح) أين ستحفظ السجلات (د) ما هى وظائف الأفراد المختلفين الذين سيعدون ويستخدمون هذه المعلومات .

(أ) محتوى السجل المجمع : أهم أنواع المعلومات فيه ما يأتى :

١ — نسبة الحضور والمعلومات المتعلقة بهذا ؟

٢ — تقديرات المدرسين للتحصيل الأكاديمى والدرجات التى أعلنت للتلاميذ أو لآبائهم .

٣ — سجلات كل الاختبارات المقننة .

٤ — سجلات متسلسلة وتقديرات غير رسمية أخرى عن التلميذ .

٥ – الفحص البدنى وبيانات عن الصحة .
٦ – سجلات عن الاشتراك فى الرياضات والأندية وأوجه النشاط خارج المنهج والعمل .
٧ – معلومات عن البيت والأسرة .
٨ – معلومات عن اهتمامات التلميذ وخططه وأهدافه .
٩ – المكانة المالية .

(ب) صياغة السجل المجمع : يجب أن نوفق بين عدة أهداف مختلفة وأحيانا متضاربة أثناء صياغة السجلات :

١ – نريد أن يتضمن السجل من المعلومات أكبر قدر ممكن .
٢ – وأن يتضمنها بطريقة يسهل قراءتها .
٣ – نريد مكانا لكثير من الفقرات .
٤ – نريد مكانا كافيا لكل فقرة منها .
٥ – نريد سجلا خفيفا يسهل ملؤه .
٦ – نريد نظاما من السجلات يمكن أن يستخدمه أشخاص مختلفون فى أماكن مختلفة .
٧ – ولكننا نريد المعلومات مركزه .
٨ – وهناك من يريد المعلومات مطولة وهناك من يريدها مقتصرة :
٩ – والبعض يريد أن تختصر السجلات حسب الطريقة المكتبية إلى أقصى حد .

والمعلومات الآتية يجب عادة أن تسجل فيما يتعلق بنتائج الاختبار المقنن :

١ – اسم الاختبار .
٢ – الصورة المعينة من الإختبار .
٣ – تاريخ تطبيق الاختبار .

٤ – الدرجة الخام .

٥ – نوع الدرجة المحولة .

٦ – الدرجة المحولة نفسها .

(ح) وضع السجلات المجمعة : وجدت السجلات لكى تستخدم ولهذا يجب أن توضع حيث يكون للجميع الفرصة المشروعة لاستخدامها والإفادة منها . فى نفس الوقت يجب أن تحمى السجلات من شر أن يستخدمها أشخاص غير مسئولين . ومشكلة المعلومات السرية تحل بأن توضع فى سجل خاص منفصل ويشار إليها فى السجل الكبير برقم وهكذا فمن يريد الرجوع إليها يستدل عنها بواسطة الرقم . وتوضع السجلات الأصلية فى مكان بحيث :

١ – يسهل أن تستخدم فى أى وقت طوال اليوم الدراسى .

٢ – يكون هناك مكان للاطلاع عليها .

٣ – أن تكون تحت إشراف شخص يستطيع أن يتحمل مسؤوليتها .

(د) وظائف أعضاء الهيئة المختلفين : يجب أن يكون نظام بطاقات التجميع عملا تعاونياً يعتمد على كل مجتمع المدرسة . فيعلم التلاميذ أن البطاقات خططت لصالحهم وأنها جزء من برنامج التوجيه . فيقدم التلميذ جزءاً من المعلومات : عن البيت ، ونشاطه خارج المدرسة ، واهتماماته ، وخططه ، وأهدافه . ومدرس الفصل يجب أن يكون مساهماً ومستخدما لبطاقات التلاميذ . فهو يساهم فيها بوضع بيانات تقييمه عن نمو الطالب ودرجاته فى المـــواد الدراسية ويستخدمها ليكون فكرة أفضل عن تلاميذه أفراداً .

جهود وزارة التربية والتعليم

قامت إدارة البحوث الفنية بوزارة التربية والتعليم بالإقليم الجنوبى سنة ١٩٥٧ بإعداد ونشر :

١ ـ البطاقة المدرسية لتلاميذ المرحلة الابتدائية .

٢ ـ البطاقة المدرسية لتلاميذ المدارس الإعدادية .

٣ ـ البطاقة المدرسية لتلاميذ المدارس الثانوية .

٤ ـ البطاقات المدرسية : وهو بمثابة كراسة تعليمات للبطاقات المدرسية الثلاث . والنواحي التي تشملها البطاقة هي :

(أ) البيانات الأولية : الاسم والسن والفرقة والفصل و . . الخ .

(ب) التحصيل المدرسي : المواد وتقديراتها في سنوات الدراسة المختلفة بهذه المرحلة من التعليم .

(جـ) المواظبة في كل السنوات : بعذر مقبول أم بدون عذر أو بعذر غير مقبول .

(د) الناحية الصحية : قوة النظر وقوة السمع وسلامة النطق والحالة الصحية العامة .

(هـ) صفات الشخصية : الجدية في العمل ، الاعتماد على النفس ، تحمل المسؤولية ، القيادة ، إبداء الرأي ، روح التعاون ، الانطواء والانبساط ، والاتزان الانفعالي ، . . الخ .

(و) النواحي الاجتماعية : تكوين الأسرة ، عدد غرف المسكن ، ووسائل شغل الفراغ .

(ز) الاختبارات : نوع الاختبار ، اسمه ، التاريخ ، الفرقة ، النتيجة .

(ح) أوجه النشاط الحر : رياضي ، فني ، ثقافي ، واجتماعي .

(ط) مشكلات التلميذ .

(ى) الملحوظات العامة .

(ك) رأى المدرسة في توجيه التلميذ .

وهكذا نرى أن هذه البطاقات هى سجلات مجمعة مصممة لمواجهة الحاجة إلى جمع بيانات عن التلميذ الفرد تتوفر فيها الخصائص التى عرضناها . إلا أنها لا تستخدم فى كثير من الحالات ومرجع عدم استخدامها أساساً إلى :

١ ــ وقت المدرس لا يتسع لهذا العمل .

٢ ــ عدد تلاميذ الفصل الواحد كبير .

٣ ــ ليس لها مكان ولا شخص مسؤول عنها فى المدرسة .

٤ ــ لا يمكن ــ حالياً ــ تغطية الكثير من بياناتها .

عرض نتائج برامج الاختبارات على الجمهور

فى عرضنا لنتائج الاختبار على الجمهور ، نهتم بعدة أنواع من المقارنات :

١ ــ مقارنة أداء الجماعة المحلية بالمعايير القومية أو بمعايير أكثر تخصصاً إن وجدت .

٣ ــ مقارنة التحصيل فى المواد المختلفة.

٢ ــ مقارنة أداء الجماعة المحلية بمستوى القدرة فى الجماعة .

٤ ــ مقارنة المدارس المختلفة فى النظام المدرسى الواحد أو بين فصول مختلفة من نفس المدرسة إن أمكن .

٥ ــ المقارنة بين الجماعات التى تعلمت بطرق مختلفة ــ التى استخدمت طرقاً مختلفة أو مواداً مختلفة .

٦ ــ المقارنة بين الجماعات فى مستويات الفرق الدراسية المختلفة .

٧ ــ مقارنة نفس الجماعة فى أوقات مختلفة لنبين نمو التلميذ .

ثانياً : هناك ثلاث صفات يحسن أن نجدها فى تفسير الدرجات للطالب أو أبيه :

١ ــ يجب أن يوضع التفسير فى الإطار المرجعى للتلميذ المعين . وعليه ،

فإن نتائج الاختبار التحصيلى المقنن يجب أن تفسر على أساس ما نعرفه عن استعداد التلميذ وعن أهدافه التعليمية والمهنية .

٢ — يجب أن يوجه التفسير نحو الفعل الإيجابى البناء . فيؤكد على خواص الصفحة النفسية للاستعداد ، أو يوجه إلى الفعل العلاجى عند ما يكون التحصيل أقل مما نتوقعه على أساس الاستعداد . فيشير التفسير هنا إلى العلاج التعليمى أو المهنى .

٣ — يجب أن يعتمد التفسير على الحقائق ولا يكون عاطفياً كما يجب ألا يبدو أحكاماً تصدر على الفرد . فيجب أن تذكر نتائج الاختبارات والأدلة الأخرى بصدق ودقة ، ولكن بطريقة ودية مقبولة .

على أنه يجب دائما أن يعلم التلاميذ نتائج الاختبارات التى طبقت عليهم . فإن الاختبار يكون ذا معنى ودلالة لدى التلاميذ إذا رأوا بعض نتائجه وفوائده واستخداماته . وأى تقرير يجب أن يكون معدلا بحسب درجة نضج الجماعة . وعلى أى حال هناك عدة طرق أخرى لتقديم هذه البيانات نذكر منها :

١ — مدرس الفصل قد يقدم تقريراً ملخصا بسيطا وتفسيراً عن أداء الجماعة ككل . ويشير فيه إلى نواحى القوة والضعف . وتقام المقارنات مع المعايير العامة لمستوى الفرقة الدراسية . ويشار فيه إلى الطريقة التى تكون بها النتائج مفيدة لعمل الجماعة وللتخطيط لأفراد معينين فى الجماعة .

٢ — ترسل الدعوة للتلاميذ لمناقشة الاختبارات مع المدرس . وفى المؤتمرات الفردية ، قد يعطى للتلاميذ صـــورة عامة عن دلالة نتائج الاختبار وكيف يرتبط بخطط التلاميذ .

٣ — وفى حالة كون نتائج الاختبارات ذات مضمونات هامة ، مثلا وجد تفاوت كبير بين الاستعداد والتحصيل ، وتفاوت كبير بين أداء الاختبار

وما يعتقده المدرس عن التلميذ من صورة ـ الذات ، وقد يوجه المدرس الدعوة إلى التلاميذ ويناقشهم فى نتائج الاختبار أفراداً وربما كان مع كل تلميذ والداه :

صعوبات برامج الاختبارات فى المدرسة

مشاكل الصدق

١ ـ تمثيل المحتوى Sampling of Content : لا يكون الاختبار عادلا فى المعرفة الحقيقة فى مجال كلى إلا إذا كان يضع أوزاناً متكافئة لكل المعالم الجوهرية فى المجال . فاختبار التاريخ ، هذا النوع من التاريخ ، هو الذى تنوى المدراس تدريسه . والاختبار ، باعتباره عينة ، قد يمثل وحسب نواحى من المواد الدراسية ، وهنا تكون هناك فرص أمام بعض التلاميذ أو بعض المدراس أن تتفوق لأنه حدث ، مصادفة ، أن كانت متفوقة فى مادة يتضمنها عدد من فقرات الاختبار .

وأحد المزايا الهامة للاختبار الموضوعى على اختبار المقال هى أن الأول يقلل احتمال أن يحصل التلميذ الضعيف على درجة عالية لأن اختبار المقالة يختار الأسئلة بطريقة مبتورة فالفرص فية قليلة . فإذا حدث مثلا أن الاختبار تضمن سؤالين يمثلان مجالين صغيرين من المجال الدراسى وحدث أن كان هذا التلميذ الضعيف يتقن الاثنين ، برغم ضعفه الشديد فى بقية المجالات الأساسية ، نال التلميذ الضعيف درجات عالية فى هذا الاختبار : بينما فى الاختبار الموضوعى ، بسبب شمول تمثيله للمجال المقاس ، يحوى فرصاً متعددة إذا عرف التلميذ الضعيف إحداها أو بعضها فلن يعرفها كلها وعلى هذا يقل احتمال حصوله على درجة لا تمثل مستواه فعلا : فإذا سألت مدرساً عن الغاية من تدريس مادة ما . فإنه سيعدد لك

أهدافاً كثيرة . فمثلا مدرس الهندسة ، قد ينكر أن الغرض من تدريس الهندسة هو توصيل عدد معين من النظريات إلى أذهان التلاميذ وتوصيل بعض الأسس التطبيقية وبعض المهارات فى استعمال الفرجار والمنقلة . قد ينكر كل ذلك ويؤكد بدلا عنه أن الغرض هو تكوين عادات فكرية ومهارات فى تحقيق الفروض ، والحذر فى تقبل استنتاجات غير محققة . ولكن لو فحصنا الاختبارات الموضوعية التى يضعها هذا المدرس نفسه فإننا نجد أنها تؤكد الأهداف الأولى التى أنكرها وتهمل الأهداف التى ذكرها .

وقد عالجت الاختبارات التحصيلية الحديثة هذا النقص الخطير بدراسة المناهج الدراسية وحصرها بدقة واعتبارها عند تصميم الاختبار عينة لها : إذ أن الاختبار ، باعتباره عينة ، يجب ، وهذا ما حدث فى الناجح منها فعلا ، يمثل المجال المقاس بكل محتواه وأبعاده بحسب نسبها وأوزانها : ذلك إما بوضع أسئلة كثيرة تغطى كل العوامل الأساسية فى المجال بحيث يخصص عدد أكبر من الفقرات للسؤال عن العوامل الأكثر أهمية . أو يخصص عدد متساوى من الفقرات للسؤال عن مختلف العوامل التى تكون المجال المقاس ، ولكن تعطى أوزان وقيم أكبر للعوامل الأهم .

ومثل هذا الاختبار سيعطى درجة منخفضة للفصول التى تركز على موضوع أو اثنين وإن تفوقت فيها تماماً . ولكن تلاميذ هذه الفصول أقل تفوقاً فى المجال الكلى كما يدرس فى أغلب المدراس :

٢ — صدق المنهج Curricular Ualidity الاختبار إذا كان عينة ممثلة لمحتوى المجال يكون اختبارا صادقاً من وجهة واحدة . وقد يكون

غير صادق من وجهة أخرى . فعندما يستخدم الاختبار للحكم على فعالية الخبرة التعليمية ، يكون صدق المنهج فيه ضرورياً . والاختبار الذى يتمتع بصدق المنهج يقيس ، بكفاءة ، المدى الذى تعلمه من الوظائف التى يهدف المنهج الدراسى أن يعلمهم إياها .

٣ – قياس جوانب السلوك المناسبة Measurement of Appropriate Behaviors لا يصح أن نقتصر على قياس كل التحصـــيل عن طريق قياس المعرفة بالحقائق العلمية أو المعلومات أو مجموعة محددة من المهارات . إذ أن المدارس تختلف من حيث عدد وأنواع المناهج وحتى لو حدد العدد والنوع فإنها تختلف أيضاً من حيث الأهمية التى تعطيها للأنواع المختلفة . هذا فضلا عن أن المدرسين أنفسهم يختلفون فى وضعهم لفلسفة وأهداف تدريس مادة معينة . وعلى أى حال فإن أهداف التعليم المدرسى متعددة جداً وقد جمعها Tyler وزملاؤه فى الآتى :

١ – المعلومات الوظيفية . أى ليست مجرد معرفة جامدة ، ولكنها معرفة يمكن تطبيقها على المواقف الجديدة التى ترتبط بها .

٢ – مهارات التفكير وعاداته .

٣ – الاتجاهات والحساسية الاجتماعية (التحمل Tolerance ، والروح العلمية فى الاستقصاء ، وتفهم الموسيقى ، و . . الخ) .

٤ – الاهتمامات والأهداف والأغراض (الأهداف المهنية ، والاستخدام الحكيم لوقت الفراغ ، . . الخ) .

٥ – دراسة المهارات وعادات العمل .

٦ – التوافق الاجتماعى والشخصى .

٧ – الإبداع

٨ ـ الصحة الجسمية .

٩ ـ فلسفة وظيفية للحياة .

وللحصول على دليل عن نمو التلميذ متقدما نحو تحصيل هذه الأهداف ، يجب أن نعرفها تعريفا إجرائيا Operational Definition خلال أنماط السلوك التى تتعلق بها أو تتضمنها .

فمثلا إن أردنا أن نقيس سلوك الطالب فيما يتعلق بقدرته على الاستنتاج من الحقائق العلمية فإننا قد نسائل أنفسنا عدة أسئلة منها أن نسأل : كيف يسلك شخص ذو قدرة عالية على الاستنتاج فى مثل هذا الموقف ؟ أو أن نسأل : كيف يمكننا التعرف على التلميذ الذى له فلسفة وظيفية فى الحياة ؟

ومن أهداف التعليم ما يسهل تحويله إلى صور سلوكية ومنها ما يصعب تحويله . ولكن إذا نجحنا فى ترجمة الأهداف الأخيرة إلى صور سلوكية سهل علينا قياسها عن طريق الاختبارات وطرق الملاحظة . فمثلا ، المهارة فى تفسير البيانات تظهر فى تصرفات خاصة . فلو أعطينا شخصا ماهراً جدولا أو رسما بيانيا فى موضوع لم يدرسه وجدنا أنه يميز الأفكار الرئيسية ويصرف نظره عن الاختلافات الجزئية التى قد تكون راجعة إلى فروق فى العينة التى أتت منها هذه الرسوم . فيرى مثلا أن العامل (أ) والعامل (ب) يتغيران معاً ولكن أحدهما لا يؤثر فى الآخر . فمثل هذه المهارات يمكن قياسها مادمنا قد استطعنا تعريفها عن طريق نواحى السلوك التى تنتج عنها .

وبعض اختبارات التحصيل تسمح لمتغيرات لا تتعلق بجوهر ما نقيسه irrelevant أن تؤثر على درجة الفرد . فقدرة الفرد على القراءة تؤثر ، غالباً ، على مستوى إجاباته عن اختبارات التحصيل . وهذه المشكلة يصعب تحاشيها ، ولكن لا يمكن الحصول ، مثلا ، على مقياس صادق للمعرفة

إذا كان الشخص الذى يعرف حقيقة معينة لا يفهم الفقرة الخاصة بها بسبب صعوبات لفظية .

هناك عامل آخر يؤثر على صدق الاختبارات التحصيلية هو عامل السرعة . والإجراء الدقيق عادة يتطلب وضع حدود زمنية ، ولكن عندما تصبح السرعة عاملا غالباً فى تحديد درجة الشخص ، يزداد احتمال كون الدرجة لا تعبر عن قدرته بدقة . والسرعة تكون عاملا مشروعا فى الاختبارات التحصيلية فقط إذا كان هدف المنهج الدراسى فعلا إعداد أشخاص يعتمدون على السرعة فى عمل معين . وهذا يصدق مثلا فى حالة إعداد مهرة فى الحساب أو موظفى خزانة ولكنه لا يصدق عند قياس صحة الاستدلال أو سعة المفردات أو المعرفة العلمية لدى المفحوص .

وفى أغلب الاختبارات التحصيلية يكون توقف الاختبار على السرعة مشروعا وحسب إذا استخدم الاختبار عمليا فى التنبؤ بالنجاح فى عمل يعتمد على السرعة ، أو إذا أيدت البيانات أن الدرجات على اختبارات السرعة ذات معامل ارتباط عالى جداً مع الدرجات على اختبار القدرة . واختبارات السرعة ذات صدق محدد عند وصف معرفة التلاميذ أفراداً ، ما دام هناك عدد قليل من التلاميذ المتفوقين يؤدى عمله ببطء .

٤ ـ التعرف والاسترجاع Recognition and Recall لأن الاختبارات الموضوعية وحدها هى الاختبارات التى يمكن أن يعتمد عليها فى مقارنة التلاميذ المختبرين فى أماكن مختلفة ، أصبح من الضرورى استخدام طريقة الاختيار من عدة إجابات وغيرها من أسئلة التعرف . كان هذا مصدراً لاهتمام مدرسين كثيرين شعروا بأن الاختبارات التى تطلب استجابات حرة وحدها هى التى نستطيع أن تقيس ، بكفاية ، ما يقومون بتدريسه ؛ ويسهل قياس التعرف على الاستجابات الصحيحة عن طريقة

الفقرات الموضوعية . ولكن المدرسين يميلون إلى تفضيل اختبارات الاستجابة — الحرة عندما يكون هدف التعليم هو إنماء القـــدرة على الاسترجاع أو اختراع حلول جديدة . وقد أيدت البحوث دائماً الاختبارات الموضوعية ذات البناء المحدد جيداً .

٥ — الدقة الفنية Technical Accuracy يتضمن الكثير من مفاتيح الاختبارات المنشورة أخطاء وهذا عيب خطير يذهب بصدقها . هذه الأخطاء تنشأ عن عدم دقة الصياغة ، والغموض ، والتبسيط المخل للمشاكل الفنية حتى تناسب صيغة الفقرة — الموصوعية . وقد جمع دياموند Diomond نقداً لـ ١٦ اختبارا نشرت عن العلوم Scienees والنظريات . وصلت نسبة الخطأ فى بعضها إلى ١٨ ٪ من الفقرات التى حوت حقائق ونظريات ليست صحيحة .

مشاكل المعايير

تستخدم الاختبارات التعليمية عادة فى تقنينها معايير الفرقة الدراسية Grade Morms ، فتحول الدرجة الخام للتلميذ إلى فرقة مكافئة grade equivalent . وعلى هذا الأساس يقال ، مثلا ، أن فلاناً من التلاميذ المتفوقين فى الفرقة السادسة تقع درجاته فى الحساب فى « الفرقة التاسعة » أى أن درجته الخام تساوى الدرجة الخام للتلميذ المتوسط فى الفرقة التاسعة فى مادة الحساب .

ولكن هذا التقرير مضلل ما لم نحط علماً بالعيوب الثلاثة الآتية :

١ — تنشأ معايير الفرقة على عينات تختلف درجة تمثيلها للتلاميذ فى الأمة بأكملها . فبعض قطاعات الأمة متفوق كثيراً على البعض الآخر ، والسبب اختلاف قدرة التلاميذ واختلاف كفاءة المدرسين ، واختلاف توسع الإقليم فى الاهتمام بالتعليم واختلاف البرامج . والأساس الوحيد العادل لمقارنة

المدارس هو الحكم على كل مدرسة مقارنة بالمدارس من نمط مشابه وذات إمكانيات مماثلة .

٢ – ليست المعايير norms مستويات standards . فمن الخطأ الشائع أن نفترض أن كل التلاميذ فى الفرقة السادسة يجب أن يصلوا إلى معيار norm – الفرقة السادسة . فنى عينة التقنين يقع ٥٠٪ من التلاميذ دون المعيار . كما أن الاختبار يجب أن يصمم لقياس المستويات الراهنة فى المدرسة وليس ما يجب أن يحصله التلاميذ تبعاً لأقصى جهد للمدارس وبأحسن طرق التدريس .

٣ – معايير الفرقة مبنية على وحدات صناعية غير متساوية من القياس . فقد يبدو أن تلميذاً معيناً ممتاز جداً لأنه وصل إلى « الفرقة التاسعة » فى الجغرافيا بينما هو فعلا فى الفرقة السادسة . وفى الكثير من معايير الاختبارات لا يعد هذا الفرق كسباً ضخماً ، إذ أن التلميذ المتوسط تنمو قدرته بمقدار صغير بين الفرقة السادسة والفرقة التاسعة .

وليس الوصول إلى مستوى « الفرقة التاسعة » صعباً بنفس الدرجة على كل تلاميذ الفرقة السادسة . فالتلميذ الذى يفوق مستوى فرقته فى مادة معينة بمقدار « فرقتين » أحياناً يعد ممتازا جدا ، وأحيانا أيضا لا يعد ممتازا جدا لأن نسبة كبيرة من فصله تستطيع الوصول ، أووصلت فعلا ، إلى ما وصل إليه فى هذه المادة . كما أن التفوق بفرقتين فى مادة معينة لا يعادل التفوق بفرقتين فى مادة أخرى . وبروفيل Profile درجات الفرق يعطى صورة جد مختلفة من حيث نقاط القوة والضعف أكثر مما يفعل بروفيل الدرجات التى تقارن تلميذ الفرقة المعينة بزملائه فى نفس الفرقة .

ويقترح كرونباخ Cronbach علاجا لذلك العيب بأن توضع معايير الفرقة الدراسية فى شكل درجات معيارية . وبعض الاختبارات المقننة الآن تستخدم الدرجات المئينية داخل كل فرقة دراسية .

خطورة برنامج الاختبارات على البرامج التعليمية

يقال إن القياس المقنن سلاح ذو حدين يحطم البرنامج التعليمى بقدر ما يساعده . والصعوبة الأساسية هى أن الاختبارات التحصيلية تحدد المنهج الدراسى . فيكرس كل من تلاميذ المعاهد والمدارس الكثير من وقتهم لمحاولة تحديد ما سيختبرون فيه حتى يتمكنوا من مذاكرة هذه المواد وإهمال ما عداها . ويستخدم التلاميذ طرقا مختلفة فى المذاكرة ويركزون على مواد مختلفة فيصلون إلى مستويات مختلفة من الإنجاز ذلك عندما تستبدل الإختبارات الموضوعية باختبارات المقال .

كما أن المدرس مدفوع بالاختبار إلى توجيه طاقاته إلى الطرق التى تمليها عليه الاختبارات . فمدرس الكيمياء مثلا يعدل منهجه ليضغط التطبيقات والاستدلال ويتوسع فى المعادلات الفنية إذا خشى أن فصله ، وبالتالى تدريسه ، سيعد فاشلا عند ما يطبق اختبار مقنن فى آخر العام تعتمد أسئلته على المعادلات . وهكذا نجد أن الاختبارات تحدد المنهج الدراسى بطريقة أو بأخرى . ولا تشـــجع الاختبارات التعليم السليم واتجاهات المذاكرة السليمة إلا عندما تقيس كل مظاهر السلوك الذى تهدف التربية إلى إنمائه .

وعند ما استخدمت الاختبارات لتحديد من فى أعضاء الجماعة أضعف من الآخرين أدت إلى الامتعاض . فبرغم زيادة الدفع ، زاد التوتر . هذا فضلا عن أن المدرسين كثيراً ما اتهموا لأن الاختبارات المقننة أظهرت أن تدريسهم ردىء . ومن العيوب أيضا أن النقل من فرقة إلى فرقة أحيانا يعتمد على نتيجة اختبار مقنن ، دون اعتبار لسجل التلميذ بالمدرسة ؛

وناظر المدرسة الذى يدخل الاختبارات يجب أن يجعلها مساعدة للمدرسين والتلاميذ معا ، وطريقة تساعدهم على معالجة نواحى الضعف . ولكن إذا لم يكن حريصا ، فقد تثير القلق الذى يخفض الروح المعنوية فى المدرسة ويقلل فعاليتها .

هناك خطأ شائع فى تطبيق الاختبارات المقننة هو استخدام المعايير دون اعتبار ما نتوقعه استدلالا من قدرة التلميذ . فما دام الذكاء يختلف من شخص إلى آخر ومن جماعة إلى أخرى فلا يمكننا أن نحكم على ما إذا كان التقدم كافيا إلا بمقارنة الشخص أو الجماعة بغيرهم ممن لهم قدرات مشابهة . وهناك طريقة شائعة تتم بها هذه المقارنة هى طريقة نسبة الإنجاز (Accompl ishment Quotient. A. Q.) ونسبة الإنجاز هى النسبة المئوية بين « العمر التعليمى » إلى العمر العقلى . فإذا كان لدينا مثلا تلميذ سنه ١١ سنة وعمره العقلى ١٢ سـنة وصل إلى معيار تحصيل تلميذ الحادية عشرة ، هذا التلميذ تكون نسبة إنجازه ٩٢ (هذا الرقم هو خارج قسمة $\frac{١١}{١٢}$ × ١٠٠)

هذه العملية لتحديد من هم « المتخلفون تحصيليا » و« المتفوقون تحصيليا » لها عيوبها التى نشأت عن تحديدات معايير الفرقة وعن الفروض الإحصائية غير السليمة هذا فضلا عن أن هذه التحديدات تفسر نسبة التحصيل وتؤيد فكرة أن المقارنة يجب أن تكون بين ما تعلمه التلميذ مقارنا بما تعلمه التلاميذ الآخرون الذين لهم استعداد أو قدرة أولية مشابهة . ومثل هذه المقارنات تستند عادة إلى المعايير المحلية . ولكن فى الإمكان إنشاء معايير على الأمة بأكملها من هذا النوع وقد قدم ترمبل وكرومباخ للمدرس الخطوات التى بها يستطيع أن يضع مثل هذا التحليل .

الإختبارات التحصيلية المستخدمة فى برامج الإختبارات

يجب أن توصلنا هذه الإختبارات إلى إجابات حاسمة عن التساؤلات الآتية :

١ – متى يجب أن تستخدم الاختبارات التحصيلية المقننة ؟

٢ – ماذا يجب أن تقيس ؟

٣ – أى الاختبارات يجب أن يختار ؟

٤ – كيف يجب أن تستخدم النتائج ؟

وأهم هذه الأسئلة هو رابعها ؛ فالوظيفة الدقيقة للاختبار فى المدرسة هى تحسين البرنامج التعليمى وهو قد يؤدى هذه الوظيفة بإحدى الطرق الآتية :

(ا) بتحديد الخبرات التى يحتاجها التلميذ .

(ب) بالإشارة إلى طرق أفضل للتدريس :

(ح) بتعديل وإنشـــاء اتجاهات لدى التلاميذ والمدرسين تسمح بتحسين التدريس .

ويقول مك كون McConn « إن كل أهداف القياس بالمدارس هى توجيه التلاميذ pupil guidance » وليس التوجيه عنده محدداً بالمهن ، بل يتسع ويصبح هادفا إلى اكتشاف كل ما يتعلق بالتلميذ حتى نوفر له نموا أفضل :

إذا اتفقنا على هذا استتبع الأمر عدة نقاط ، فإذا كانت الاختبارات بهدف الإرشاد أصبحت جزءاً من العملية التعليمية ووجب أن تكون طبيعية لا صناعية ، ولهذا كان اختيار وقت القياس أمراً هاما ، فالقياس بعد وقته

الأنسب لا يفيد كثيرا ولهذا تلجأ مدارس كثيرة إلى تطبيق الاختبارات التحصيلية المسيحية فى أول العام الدراسى .

وعند ما تتطبق اختبارات مناسبة وتوضع النتائج فى يدى كل مدرس ، تصبح أساساً سليما لتخطيط أعمال السنة . وهكذا يصبح من الضرورى أيضاً أن يعاد الاختبار فى نهاية العام الدراسى لتقدير مدى التقدم الذى أحرزه التلاميذ . ويهمنا أن يكون الاختبار تشخيصياً توجيهياً ولا يكفى أن يقدم لنا درجات فقط ، وإذا كانت الاختبارات مستخدمة بغرض التوجيه فهى إذن تستخدم للتلميذ وليس عليه . فتصبح وسيلة تعينه على معالجة نقاط ضعفه ، وتصبح سلاحا فعالا له يقنعه بالاشتراك فى برامج معينة أو تغيير عادات معينة والاختبارات هنا تقنعه أفضل من ضغط المدرس عليه . وفى مثل هذا القياس يجب أن يأخذ التلميذ الاختبارات لأنه يريد أن يعرف النتائج .

ويجب أن نهتم أكثر بنواحى معينة من الأداء — وبالنجاح فى فقرات معينة ، أو مهارات فرعية بمعينة فى الاختبارات المركبة . ويجب أن نهتم بنوع الخطأ وليس بمقداره :

كان استخدام الاختبارات التحصيلية فيما مضى على غير هدى لأنها أجبرت المدارس على تدريب التلاميذ لا تعليمهم . ولكن إذا أعطيت الاختبارات على ضوء فكرة كاملة عن ماضى التلميذ ، وإذ أريد بالاختبارات أن ترشده فى تقدمه المقبل ، فإنها ستكون خير معين .

الاختبارات النفسية فى البرامج

من العرض السابق يتضح لنا أن الاختبارات النفسية والتربوية معا هى وحدات برامج الاختبارات فى المدرسة ، وفى البرامج المقترحة يلاحظ القارئ أن الاختبارات التى حوتها تقيس وظائف ، سواء كانت تربوية

أو نفسية ، تتسق مع مراحل نمو التلميذ باعتباره فردا أولا ثم مستفيدا من برامج محددة مدرسية ثانيا . ومما يدعم استخدام الاختبارات النفسية فى المدارس أن :

١ ــ التلميذ فرد وهو مفحوص كغيره ممن يجرى قياسهم :

٢ ــ أغلب جماعات التقنين حوت تلاميذاً .

٣ ــ أننا عند ما نقيس التلميذ لا نحتاج وحسب إلى تقدير أثر التدريب الرسمى الذى ناله فى المدرسة ، بل وأيضاً إلى تقدير :

(ا) عضويته فى جماعات المدرسة (وأداة هذا القياس هى مقاييس العلاقات الاجتماعية والذكاء الاجتماعى والنضج الاجتماعى والسوسيوجرام)

(ب) صحته النفسية وتوافقه (وأداة هذا القياس هى اختبارات الشخصية ومقاييس الصحة النفسية)

٤ ــ أن نمو التلميذ يلزمنا بقياس :

(ا) قياس الذكاء والمهارات الأساسية والإستعدادات فى سن المدرسة الابتدائية والإعدادية .

(ب) قياس الاستعدادات الخاصة والميول والشخصية والتوافق فى سن المدرسة الثانوية .

(حـ) قياس الميول والتوافق فى سن المعاهد العليا والجامعة .

٥ ــ أن استعدادات تلميذ المدرسة الابتدائية تتضح فى قدراته فى نهاية سن المدرسة الإعدادية تقريباً . ومن هنا وجب قياس أثر التلمذة فيهما على هذا النمو وأثر هذا النمو على النجاح فى المدرستين :

٦ ــ أن هذه القدرات تتخصص كلما تقدم التلميذ فى التعلم وتتضح الاستعدادات الخاصة والفارقية و . . الخ مما يدفعنا إلى قياس أثر الحياة فى المدرسة على النمو وأثر النمو على النجاح المدرسى .

٧ ـ أن المدارس تعد الأفراد للمهن وهنا نستخدم اختبارات الكفاية واختبارات الحرف لقياس أثر التدريب الفنى على كفاية الفرد :

٨ ـ أن المدارس تعد الأفراد للحياة ، وإذا أضفنا إلى هذا أن أثر خبرات الطفولة والمراهقة على التكوين النفسى للفرد وتوافقه المقبل يزيد عن أثر الخبرات التالية كثيراً ، رأينا ضرورة قياس التوافق والشخصية فى مجال المدرسة .

ولهذه الأسباب يتعين علينا ، بعد دراسة إنشاء الاختبارات الموضوعية واختبارات المقـال ، أن نعرض لخصائص إختبارات التحصيل فى الفصل القادم :

الفصل الخامس عشر

خصائص ومشكلات الاختبارات التحصيلية

- خصائصها .
- استخدامها .
- تحديداتها .
- صـــدقها .
- معايير ها .
- وظائفها .
- أنواعها .

الفصل الخامس عشر

خصائص ومشكلات الاختبارات التحصيلية

خصائصها :

تعتبر الاختبارات التحصيلية أكثر أنواع الاختبارات المقننة شيوعا ؛ والهدف الأساسى للاختبارات التحصيلية هو إظهار أثر الدراسة أو التدريب الخاص . وبالرغم من أنه يغلب تطبيقها فى المدارس إلا أنها تستخدم فى ميادين أخرى ، كأن تقيس التدريب المهنى والتخصص والخبرة فى كثير من الأعمال .

ومن المعتاد أن نقارن أيضاً بين الاختبارات التحصيلية واختبارات الاستعدادات . وتختلف الاختبارات التحصيلية عن اختبارات الاستعدادات فى درجة توحيد الخبرة السابقة المقاسة . إذ تقيس الاختبارات التحصيلية أثر مجموعات من الخبرات المقننة هى المناهج الدراسية . بينما نجد أن اختبارات الاستعدادات تقيس الأثر المتجمع للخبرات المتعددة فى الحياة اليومية . ويمكننا القول بأن اختبارات الاستعدادات تقيس أثر التعلم تحت ظروف غير معروفة وغير مضبوطة بينما تقيس الاختبارات التحصيلية أثر التعلم تحت ظروف معروفة ومضبوطة إلى حد ما .

واختبارات التحصيل تختلف أيضاً عن اختبارات الاستعدادات من حيث استخدام كل منها . فاختبارات الاستعدادات تفيد فى التنبؤ بنوع الأداء المنتظر ، فهى تستخدم فى تقدير المدى الذى سيفيد به الفرد من التدريب كما تستخدم فى التنبؤ بنوع التحصيل فى موقف جديد . فى حين نجد أن الاختبارات التحصيلية تبين المكانة الحالية للفرد ومدى إفادته من

التدريب إذ يدلنا الاختبار على ما يمكن أن يعمله الفرد فى زمن محدود بعد قياس كفايته أو تحصيله فى زمن محدد .

ولكن لا يمكن أن نميز بين هذين النوعين من الاختبارات بصورة قاطعة فبعض اختبارات الاستعدادات تعتمد على تدريب خاص سابق موحد بين الأفراد ، وهنا تقترب كثيراً من الاختبارات التحصيلية . وكذلك فإن بعض الاختبارات التحصيلية تغطى مدى واسعا من الخبرات التعليمية غير المقننة وهنا تقترب كثيراً من اختبارات الاستعدادات التى تقيس أثر خبرات غير مقننة . وبالمثل فإن أى اختبار تحصيلى يمكن أن يستخدم فى التنبؤ بالتعلم المستقبل ، هذا فضلا عن استخدامه فى ميادين التحصيل الأخرى : فمثلا التقدم الذى أحرزه طالب فى مادة معينة ، حسبما تدل درجاته فى اختبار تحصيل فى المدرسة الإعدادية ، دليل نعتمد عليه فى التنبؤ بنجاح نفس الفرد فى هذه المادة فى المعاهد العليا . وعلى هذا فدرجات الأفراد فى الاختبارات التحصيلية التى تغطى المنهج الدراسى ، تفيدنا فى التنبؤ بنجاحهم مستقبلا :

وفى التميز بين اختبارات الاستعدادات والاختبارات التحصيلية يجب أن نقول ببساطة أن الاختبارات التحصيلية تقيس ، أثر التعلم المقنن وقد أصبح من الواضح أن كل الاختبارات النفسية والتربوية تقيس سلوك الفرد الحالى الذى يعكس أثر تعلمه السابق فى مختلف المجالات وقياس أثر التعلم السابق يصلح للتنبؤ بالتعلم المستقبل .

استخدامها :

تستخدم الاختبارات التحصيلية فى التعليم والخدمات المختلفة والجيش كما أنها تستخدم أيضاً فى مجال علم النفس الإكلينيكى وهى تخدم أغراضا عديدة أهمها :

١ – الاختيار أو الانتقاء Selection : كانتقاء الصالحين لنوع من التعليم، أو المدارس الفنية . . الخ .

٢ – التصنيف Classification : كتصنيف الجنود بعد معرفة خبراتهم وتدريبهم العملى السابق أو تصنيف أطفال المدارس الابتدائية والإعدادية حتى يكونوا جماعات متناسقة يضم كل جماعة منها فصل أو جماعة مدرسية .

٣ – فى الحكم على مستوى مهارات الفرد ومعرفته الحالية وهذا أمر لازم لوضع خطة تعليمية أو تدريب مهنى حتى يمكن الحكم عليه .

٤ – التحقق مما إذا كان الفرد قد وصل أم لم يصل إلى مستوى معين كما فى حالة استخدامها فى الجيش أو فى الصناعة .

٥ – فى الأغراض الإكلينيكية ، كتشخيص قدرة الفرد أو عجزه ومدى هذا العجز فى ناحية ما مثل صعوبات القراءة أو أى ضعف تعليمى آخر ومعرفة الأفراد المتمردين والجانحين . خاصة وأننا أحيانا نجد الفشل المدرسى وسوء التوافق الدراسى عند بعض الأطفال الموهوبين .

٦ – تقييم البرامج الدراسية وكذلك تعديلها ، فمن طريق الاختبارات التحصيلية نقيس فعالية البرامج ومدى تحقيقها لأهداف التربية وما ينبغى إدخاله على تلك البرامج من تعديلات .

٧ – تقييم التلميذ ، لأن موضوعية الاختبارات التحصيلية وشيوعها يساعدان المدرس على وضع تقييم أدق للتلميذ .

٨ – تعيين البرامج العلاجية (الهوايات مثلا) فهى تعين فى تحديد نوع البرامج العلاجية الملائمة للفرد ، وتحدد درجة استفادته منها حاليا ومستقبلا .

٩ – وعلى ضوء ما سبق تساعد الاختبارات التحصيلية وتيسر عمليات

التعلم وذلك بمعرفة نقاط الضعف فى التعلم السابق وتحديد طريقة الدراسة والمذاكرة التى يتبعها الطالب حاليا فتدفعه إلى التحصيل الدراسى والتعلم بطريقة أفضل : .

وقد ظهرت فائدة استخدام الاختبارات التحصيلية وأثرها فى رفع ودفع التحصيل والتعلم فى تجارب عديدة نذكر منها المثال التالى الذى ضربته أنستازى :

أخذت ستة (٦) فصول كجماعة ضابطة وأعطيت اختبارا فى أول البرنامج الدراسى واختبارا آخر فى منتصفه واختبارا ثالثا فى نهايته . وأما الجماعتين التجريبيتين فكانت كل جماعة تكون من ٤ فصول أعطيت ١٢ اختبارا أثناء الدراسة . وفى إحدى الجماعتين التجريبيتين استخدم جهاز كانت إجابات المفحوص عن فقراته تسجل بالضغط على أحد أزرار خمسة فإذا كانت الإجابة صحيحة ترك الزرار علامة حمراء وأظهر له الجهاز الفقرة التالية أما إذا ضغط على زرار يقابل إجابة خاطئة فإن الفقرة تظل أمامه ولا تظهر الفقرة التالية وهنا يحاول المفحوص الأزرار الأخرى حتى يصل إلى الإجابة الصحيحة ، وتظهر له الفقرة التالية وهكذا يعرف المفحوص أخطاءه . والدرجة على الاختبار تقل تبعا لعدد محاولات الضغط على الأزرار أما المجموعة التجريبية الأخرى فكانت نتيجة الاختبار الكلى تعلن لأفرادها فور انتهائهم منه .

وفى نهاية مدة التدريب ظهر أن درجات تلاميذ المجموعتين التجريبيتين تزيد كثيراً عن درجات تلاميذ المجموعة الضابطة . وزادت درجات تلاميذ المجموعة التجريبية الثانية ، التى كان المفحوص فيها يعرف بنفسه أخطاءه وصوابه فى كل فقرة ، عن درجات تلاميذ المجموعة الأولى ، التى كان المفحوص فيها يعرف النتيجة النهائية للاختبار الذى أجرى عليه ، زيادة كبيرة .

١٠ — وتفيد الاختبارات التحصيلية عموما فى تحسين طرق التعلم ومراجعة محتويات المناهج والوقوف على كفاية كل محتوى والمواقف التى تتطلب توحيد التدريب ومدى ونوع الأخطاء الشائعة فى فهم موضوع معين .

تحديداتها :

إننا بدون شك نعترف بفوائد الاختبارات التحصيلية غير أن البعض يغالى فى هذه الفوائد ويظهر هذا عند التطبيق . ومن أكثر الأخطاء شيوعا تقنين التعليمات فإذا كان التقنين مصحوبا بتعلم المهارات الأولية كما فى بعض المواقف المهنية والعسكرية التى قنن فيها الاختبار التحصيلى إلا أن التحصيل لا يقوم على تعلم نفس المهارات التى قنن عليها فى مجالات أخرى من التدريب . وعلى أى حال فإن اضطراد توحيد المهارات يقضى على التلقائية والإبداع والأصالة فى التفكير ، وأكثر من هذا فإن التقنين يتجاهل الفروق الفردية فى المعلم والمتعلم ويتجاهل الحاجات والظروف المحلية الخاصة لكل من الفاحص والمفحوص .

وما زالت أغلب الاختبارات التحصيلية تهتم ببعض التعليمات أكثر من غيرها ، فمن السهل مثلا أن نضع أسئلة تتعلق بكل تفاصيل المنهج الدراسى المعين بينما يصعب ذلك فى حالة السؤال عن الفهم والإدراك النقدى وتطبيق المبادئ العامة . إلا أنه يغلب نسيان التفاصيل الواقعية المتعلمة خلال الدراسة بينما نجد تفسير المبادئ المتعلمة وتطبيقها فى مواقف جديدة يمثل القدرة على الاسترجاع والكسب الحقيقى بعد مرور زمن من إتمام المنهج ، ولهذا يعاب على الاختبارات التحصيلية أنها غالباً ما تشحن بمواد لها قيمتها الحالية فى الدراسة ، وتقل قيمتها فيما بعد عند مواجهة الحياة الواقعية .

وبالرغم من كل هذه التحديدات للاختبارات التحصيلية فإن التلاميذ والقائمين بالاختبارات يكونون بفضلها مدفوعين دائماً إلى الاهتمام بالتركيز على المواد التى ترفع درجاتهم فى هذه الاختبارات .

ويجب أن ينظر إلى الاختبارات التحصيلية على أنها أدوات وليست أهدافا . وأكثر من هذا فعلينا أن نذكر أنها تعطينا بيانات جزئية ليست كافية لفهم الفرد ولهذا نحتاج إلى ملاحظات أخرى إضافية . وهذا وجه من أوجه الشبه بينها وغيرها من أنواع الاختبارات الأخرى . والواقع أن الاختبارات التحصيلية فى تحسن دائماً فهى الآن تقيس مظاهر التعلم المعقدة. وتتنبأ من خلال هذه الدرجات بالنجاح فى الدراسة أو المهنة مستقبلا . إلا أن بعض ميادين التربية يصعب فيها تصميم اختبار جيد . ومن أهم هذه الميادين قياس القدرة الناقدة ، وقياس استخدام المعارف القديمة فى مواقف جديدة . ومع هـــذا فالاختبارات التحصيلية تخطو الآن سريعاً نحو هذه الميادين :

صدق الاختبارات التحصيلية

لا تختلف طرق حساب معاملات الصدق ومعانيه فى الاختبارات التحصيلية عنها فيما عداها من الاختبارات ، إلا من حيث تفضيل استخدام طرق معينة تتناسب مع طبيعة المادة ، واستخدام محكات معينة تتوفر فى أكثر من غيرها فى هذا المجال . ففى الاختبارات التحصيلية يوجد الصدق بطريقتى تحليل المحتوى Content analysis والتحليل المنطقى logicol analysis . وهذا لا يعنى أنهما الطريقتان الوحيدتان فى ميدان التربية إذ تستخدم كل الطرق الأخرى فى إيجاد الصدق .

وأهم أسباب شيوع استخدام طريقتى التحليل المنطقى وتحليل المحتوى أن واضع الاختبار التحصيلى لا يجد وقتاً كافياً لدراسات مستفيضة إذ أن الاختبار التحصيلى الجيد هو الاختبار الذى يساير المنهج الدراسى الذى يقيسه . وهذا المنهج يختلف من بيئة إلى أخرى ومن نوع معين أو مرحلة معينة من التعليم إلى آخر ، ناهيك عن اختلافات بعض المناهج بين

مدارس المنطقة الواحدة واختلافات طرق تدريس المقرر الواحد فى المادة المعينة من مدرس إلى آخر ومن فصل دراسى ذى مستوى معين إلى آخر ؛ كل هذا يوجب عمل اختبارات تتحدد بمجال معين ضيق ، وبالتالى تتعدد اختبارات التحصيل التى تقيس نفس المادة فى نفس الفرقة الدراسية . وهذا التعدد من شأنه أن يجعل اختبارات التحصيل عرضة دائماً للتغيير والتجديد مما لايسمح بفرصة كافية لقيام الباحث بدراسات الصدق والثبات وغيرها من البحوث التى تلتزم بها الاختبارات فى المجالات الأخرى .

فإذا أضفنا إلى هذا أن طبيعة المادة المقاسة ، وهى غالباً مقرر موضوع معين ، تحول بين الباحث وبين القيام بدراسات تجريبية تتبعية وبينه وبين القيام بأنواع أخرى من التحليل مثل التحليل العاملى ، وجدنا أن أنسب طرق إيجاد الصدق فى الاختبارات التحصيلية هما طريقتى تحليل المنهج والتحليل المنطقى .

وفى تحليل المنهج يقوم واضع الاختبار بدراسة الأهداف ودراسة المحتوى أى المنهج الدراسى وتحديد أجزائه ونسبها . ثم يقوم بوضع فقرات تناسب كل هذا ، واضعا فى اعتباره أن يمثل عدد من الفقرات جزءاً من المحتوى بحسب وزن هذا الجزء بالنسبة للمنهج المقاس كله . أما فى التحليل المنطقى فيقوم الباحث بوضع المفاهيم التى يود أن يكون التلاميذ قد حصلوها بعد أن درسوا المنهج ، ثم يعالج الفقرات منطقيا لتبين درجة تغطيتها لهـذه المفاهيم متمسكا بتحصيل الأهداف الذى يميز بين التلميذ الضعيف والتلميذ المتفوق . وفى الحالتين يهتم الباحث بأن يكون الاختبار كافيا فى تمثيل المنهج الدراسى .

والمحك الذى يغلب استخدامه فى الاختبارات التربوية هو النجاح الدراسى . فإذا كان أغلب الحاصلين على درجات مرتفعة فى الاختبار ناجحين دراسياً ،

كان الاختبار صادقاً . وإذا تزايدت نسبة الناجحين فى الفقرة بارتفاع مستواهم ،كان الاختبار صادقا . وهنا نختار الفقرات التى تميز بين الأفراد بما يتفق مع تميزهم إلى مستويات دراسية مختلفة . ولكن استخدام تقدم الدرجات إذا ارتبط بالنجاح الدراسى ، وتزايد نسب الناجحين فى الفقرة إذا تدرجت مستوياتهم الأكاديمية ، استخدام هذين الدليلين كمحك فى الاختبارات التحصيلية ليس مشروعا دائما .

وأول أسباب عدم كفاية هذا المحك ، الاختلافات البيئية والإقليمية بين محتويات ومستويات المنهج الدراسى ، وقد سبق الإشارة إليه . كما أن بعض المواد الدراسية يتوقف تدريسها فى مراحل معينة ، وهنا لا يحصل تلميذ الفرقة الأعلى على درجة أكبر بل وقد يحصل على درجة أقل من تلك التى يحصل عليها من يدرس هذه المادة . وليس هذا ، بالطبع ، دليل على انخفاض مستوى تحصيل طالب الفرقة الأعلى ، إذ أن طالب الفرقة الأقل لو اختبر بنفس الاختبار بعد انتهاء تدريسه هذه المادة بمدة ، لحصل على درجة أقل من درجته الأولى .

وأخيراً ، فإن الاختبارات التحصيلية قد تستخدم فى التنبؤ بالأداء المقبل بافتراض أنه متسق مع الأداء الحالى الذى تقيسه . وفى هذا الغرض يجب بحث الصدق التجريبى بالطريقة التتبعية ، وهنا ينهج الأخصائى النفسى التربوى منهج الأخصائى النفسى فى مجالات القياس الأخرى . (انظر الفصل الرابع عشر) .

معايير الاختبارات التحصيلية

المعيار أساس عام مشترك به تصبح الدرجة الخام ذات معنى فتفسر درجة الفرد بمقارنتها بنقطة تنسب إليها . وبينما يطبق هذا المفهوم فى

اختبارات الذكاء باستخدام معيار العمر العقلى ، نجده يطبق فى الاختبارات التحصيلية باستخدام معيار الفرقة الدراسية Grade Norm .

وسبب استخدام معيار الفرقة الدراسية هو أن التلاميذ مقسمون إلى جماعات متجانسة مكونة بالفعل هى جماعات الفصول والفرق . فلا يحتاج الباحث إلى القيام بتصنيف المفحوصين إلى جماعات يحسب لكل منها معيار يناسبها ؛ كما أن معيار الفرقة الدراسية واضح يمكن تفسيره مباشرة . ونحصل على هذا المعيار بحساب متوسط درجة كل فرقة دراسية بالنسبة للمادة التى يقيسها الاختبار . وتحول الدرجة الخام للفرد إلى مكافئ الفرقة Grade Equivalent إذا كانت تساوى متوسط درجات تلاميذ هذه الفرقة ؛ وهكذا يصبح التلميذ فى مستوى فرقة متقدمة فى إحدى المواد لأنه يتفوق فيها ، وفى مستوى فرقة متأخرة فى مادة أخرى لأنه ضعيف فى هذه المادة .

وبهذه الطريقة نعتبر أن التلميذ فى عدة فرق دراسية ما دامت درجاته فى المواد المختلفة تصل إلى مستوى هذه الفرق ، أى تساوى متوسط درجات التلاميذ الذين أخذناهم من هذه الفرق فى عينات تقنين تلك الاختبارات . وهذا يمكننا من رسم صفحة نفسية للتلميذ (انظر شكل (١١) الفصل التاسع) .

وقد اقترح كيلى T.L. Kelly معيار العمر المنوالى modal-age ، بمعرفة العمر الشائع للفرقة الدراسية المعينة واستبعاد من يزيدون عنه أو يقلون . ثم يحسب متوسط درجات أفراد هذه الفرقة الدراسية بعد أن أصبحوا من عمر واحد . وتتراوح أعمار أفراد الفرقة الدراسية بهذه الطريقة فى مدى سنة كاملة . وما دام مدى كبر أو صغر سن التلميذ عن منوال فرقته يختلف باختلاف المدارس وباختلاف الفرق الدراسية داخل المدرسة

الواحدة ، فإن المعيار الذى يحسب من درجات مجموعة أفراد الفرقة الدراسية لا يمكننا من مقارنة الفرق الدراسية المختلفة .

وبرغم مزايا معايير الفرق الدراسية وسهولتها العملية ، إلا أن ثمة اعتراضات كثيرة تقف دونها . وأول هذه الاعتراضات هو أن محتوى ومستوى التعليم يختلف من فرقة إلى أخرى . كما أن حصول التلميذ على درجة تعادل متوسط فرقة دراسية أعلى من فرقته لا يعنى وصوله فعلا إلى مستواها فى مادة الاختبار نتيجة لتحصيله الرسمى . بل قد يعود ارتفاع الدرجة إلى القراءات والاطلاع ، وهـذا يعنى اتفاق المادة مع ميول الطالب واهتماماته بالوظائف المقاسة فى الفرقتين .

والنقد الذى وجهناه إلى مفهوم العمر العقلى ينطبق أيضاً على مفهوم معيار الفرقة ونقصد أن الوحدات غير متساوية . بمعنى أن نمو الفرد فى الأعمار الصغيرة يكون بسرعة أكبر منه فى الأعمار الكبيرة ، والفرق بين مستوى الفرقة الدراسية الثانية والثالثة قد يكون أكبر من الفرق بين مستوى الفرقة الثالثة والرابعة إذ أن قدرة التلميذ على التحصيل فى الفرق الأدنى تزيد عنها فى الفرق الأعلى .

ومن عيوب معيار الفرقة أيضاً ، أنه خلال أى فرقة دراسية واحدة ينمو التلاميذ فى معرفة المواد بسرعات تختلف من مادة إلى أخرى بحسب المنهج الدراسى . فالتلميذ الذى يعلو فرقته الدراسية بفرقتين فى مادة معينة ويعلوها بفرقتين أيضاً فى مادة أخرى قد يزيد عن متوسط فرقته فى المادة الثانية بدرجة أكبر من زيادته عن متوسط الفرقة فى المادة الأولى لأن الفرقة فى المادة الأولى تتقدم علمياً بسرعة أقل منها فى المادة الثانية . وفى هذه الحالة نجد أن بعض تلاميذ هذه الفرقة يفوقها بفرقتين ولا يلزم أن قدراته فيها عالية .

وتدريس مقررات خاصة فى مادة معينة قد يجعل الفرد متفوقاً فى هذه المادة ولكن تفوقه لا يصل إلى مستوى فرقة دراسية أعلى من فرقته ؛ وإذا كان الاختبار التحصيلى مصمما لقياس مدى تفوق التلميذ بالنسبة لفرقته ، وحصل التلميذ على درجة عالية ، كان من الخطأ أن نعتبر أنه وصل إلى مستوى فرقة دراسية أعلى لأن الاختبار يقيس التفوق داخل الفرقة بالنسبة لمجموعتها ولا يقيس الوصول إلى مستويات فرق أعلى من هذه الفرقة .

فإذا حصل تلميذ فى السنة الأولى الإعدادية عن اختبار حساب على الدرجة ٣٥ . وكان متوسط درجات تلاميذ السنة الثالثة الإعدادية فى نفس الاختبار هو ٣٥ ، فيجب أن نعتبر هذا التلميذ متفوقا فى فرقته . ولا يصح اعتباره فى مستوى الفرقة الثالثة ، وإلا أصبح معيار الفرقة مضللا إذ يحصل التلميذ المتفوق فى السنة الأولى على درجة التلميذ المتوسط فى الثالثة ؛ هذا بينما تختلف فى المادة الثانية وأن عدداً أصغر يفوقها بفرقتين أيضاً ؛

وآخر عيب نذكره ، لمعيار الفرقة الدراسية هو أنها لا تشمل أفراداً متجانسين بدرجة كبيرة فهناك فروق فردية بين تلاميذ الفرقة الواحدة تجعل بعضهم يفوقها بعدة فرق والبعض الآخر يقل عنها بعدة فرق خاصة فى المواد التى يرتبط فيها التحصيل بالميل أو الاهتمام أو يتأثر فيها الفرد بخبراته أو أسرته . وبناءاً عليه لا يصح اعتبار الفرقة الدراسية كلها من مستوى أداء واحد فهناك فروق فردية قد تكون واسعة المدى أحياناً ؛

وعلاجا لكل هذه العيوب لجأت الاختبارات التحصيلية إلى أنواع أخرى من المعايير وإلى تعديلات على معيار الفرقة الدراسية كان أولها معيار العمر المنوالى . وقد تعرضنا فى الفصل التاسع والفصل الرابع عشر لأغلب هذه المعايير وهى :

١ ـ معايير العمر التعليمى .

٢ ـ النسبة التعليمية ؛

٣ ـ نسبة التحصيل أو الإنجاز .

٤ ـ المئين داخل الفرقة الدراسية .

وقد بدأت الاختبارات التحصيلية تعتمد على معيار تستخدم فيه توزيعاً مركبا واحداً لكل الدرجات المعيارية فى كل الفرق الدراسية ؛ فتختار فرقة دراسية معينة كجماعة مرجعية ، ويستخدم متوسط هذه الفرقة وانحرافها المعيارى فى حساب الدرجات المعيارية . فبغض النظر عن العمر أو الفرقة الدراسية تحول الدرجة الخام إلى مكافئها الدرجة المعيارية ؛ ومن هنا يفيد هذا المعيار فى المقارنة وفى رسم وتفسير الصفحات النفسية ؛

ويرى البعض عمل توزيع بحسب السن ، ويرى آخرون عمل توزيع بحسب الفرق الدراسية المقاسة ذلك داخل الفرقة المرجعية باختبار يعتمد على هذا النوع من الدرجات المعيارية الكلية . وهكذا تحول الدرجة الخام للتلميذ إلى درجة معيارية عالية إذا كان سنه صغيراً أو من فرقة دراسية منخفضة ، ويحصل التلميذ الأكبر عمراً أو من فرقة دراسية أعلى على درجة معيارية كلية أقل لأن عمره أو فرقته تسمح له بتحصيل أفضل . بينما يحسب البعض الآخر المقابلات المئينية لهذه الدرجة المعيارية الكلية فيعد التلميذ فى المئين ٧٠ مثلا لأن درجته المعيارية تقابل هذا المئين ؛ ولكن كل هذه التعديلات تواجه النقد الذى سبق أن وجه إلى معايير العمر والمعايير المئينية ومعايير الفرقة الدراسية ؛ وأحدث المعايير استخداما فى اختبارات التحصيل يعتمد على تحديد مستوى يفترض أن التلميذ وصل إليه ما دام قد أنهى برنامجا دراسيا معينا ؛

وظائف الاختبارات التحصيلية

وقد يحس القارئ تداخلا بين الاستخدامات Uses والوظائف Functions ونحن نوافقه إلى حد ما . وبيان علاقة الجزء بالكل بين الاستخدامات والوظائف يقتضي مناقشة فلسفية ليس هذا مجالها . وأفضل طريقة لحسم هذا النقاش هي عرض تقرير مكتب شئون الأفراد بالبحرية الأمريكية . وفي هذا التقرير عرض للأغراض التي نجحت في تحقيقها الاختبارات التحصيلية المقننة وأهمها :

١ ـ ساعدت الاختبارات في تقنين التعليم لكل المدارس التي تعد الناس لنوع معين من الوظائف . فبالرغم من أن أهداف التعليم والمناهج كانت مقننة ، إلا أن بعض المدارس قد انحرفت بإهمال مواضيع معينة أو المغالاة في الاهتمام بها . وأي انحراف سينعكس في الدرجات على الاختبار فتنخفض درجات المدرسة عن درجات بقية المدارس والاختبارات لهذا أجبرت المدرسين على أن يسيروا حسب ما اعتزم واضعو المنهج .

٢ ـ تعتبر الاختبارات أساساً يوحي بمراجعة المنهج الدراسي وتحسين طرق التعليم . فإذا وجد من الاختبار أن الوقت لا يسمح بإتقان مواد دراسية معينة ، كان العلاج هو إعطاء وقت إضافي في المنهج لهذه المواد . فقد كان المدرسون ، قبل تطوير الاختبارات التحصيلية ، يعتقدون أن إنهاء جزء من المنهج يعني أنهم درّسوه على خير وجه . ولكن نتائج الاختبارات أصبحت معيناً للمدرسين يلجئهم إلى المشرفين والمتخصصين ليحسنوا طرقهم في التدريس . وكأن الاختبارات ساعدت المدرسين على أن ينموا نقد ـ الذات Self- Criticism

٣ ـ تطلبت بعض الاختبارات أن يظهر التلاميذ مهاراتهم في العمل أكثر من مجرد الإجابات اللفظية . وهكذا وجهت الاختبارات الانتباه إلى

أهمية تعليم السلوك الذى يهدف المنهج إلى إنمائه لدى التلاميذ . فانخفضت أهمية المناقشة فى المحاضرة كطريقة فى التدريس . ووجد أن فعالية التدريب قد زادت .

٤ ـ كانت درجات التلاميذ تعطى على أساس الانطباع الذاتى فى حالة عدم وجود الاختبارات التحصيلية أو ، على الأقل ، فى حالة عدم إجراء اختبارات يضعها المدرسون . ومثل هذه الدرجات لا يعتمد عليها ، وقد فشلت فى لفت النظر إلى كل الجوانب الجوهرية من المنهج ، وخضعت للتعصبات . والاختبارات الموضوعية ، نفسها ، عندما يضعها مدرسون لم يأخذوا تدريباً خاصاً ، كانت دائماً أدوات ذات درجة منخفضة من الموضوعية . والإعداد الدقيق للاختبارات يزيد احتمال حصول كل فرد على درجته بدقة تامة .

٥ ـ الاختبارات التى أكدت على كل الجوانب الجوهرية من المقرر أثبتت أن التحصيل بصفة عامة أصبح موضع اعتبار وأن النقائص فيه أصبحت ملحوظة ففى كلية الهندسة مثلا ، قبل الأخذ بالاختبارات المقننة ، كانت الدرجات تتأثر كثيراً بأداء النواحى الرياضية من المنهج ، وربما كان هذا لأن هذه القدرة يسهل قياسها . فتغيرت الامتحانات لتضع التأكيد على الفهم الميكانيكى . ونتيجة لهذا ، وجه الانتباه إلى الأشخاص الذين برغم مهاراتهم الحسابية ضعاف فى مواد أساسية أخرى .

٦ ـ سمحت الاختبارات المقننة بتحليل الدرجات ومستوياتها وقللت الفروق الراجعة إلى كرم « بجبحة » أو قسوة المصححين .

٧ ـ تحسن الدافع لدى التلاميذ والمعلمين بقيام المنافسة على أساس عادل . فأحسن طريقة لدفع الدارس هى أن نوضح له الطرق المعينة التى يحتاج فيها إلى إنماء مهاراته .

٨ – فى إنشاء اختبارات انتقاء التلاميذ أصبحت الدرجات النهائية الدقيقة محكاً أفضل على أساسه يشتق صدق هذه الاختبارات .

٩ – حطمت الاختبارات المقننة اعتبار طول مدة التعليم مقياساً لقدرة المتعلم . فقد وجد أنه لا يمكن اعتبار عدد السنوات ولا عدد المناهج أساساً للحكم على الشخص . ففى دراسة أجريت سنة ١٩٣٨ أخذ آلاف الطلبة الجامعيون اختبارات مقننة للمعرفة فى ميادين متعددة ؛ ووجد أن كثيراً من طلبة السنوات النهائية فى الجامعة يعرفون أقل مما يعرف التلميذ المتوسط فى السنة النهائية بالمدرسة الثانوية .

١٠ – وللاختبارات التحصيلية أيضا فوائدها الإكلينيكية والتنبؤية ، فضلا عن وظيفتها العامة فى تقييم التعليم . واختبارات التحصـــيل غالباً ما تكون أداة ممتازة فى التنبؤ بنجاح التلميذ باختيار مناهج أعلى من التى أخذها التلميذ فعلا . وهى قد تستخدم فى اختيار الموظفين – كما أن الفوائد الإكلينيكية لاختبارات التحصيل تتضح فى حالة الجانح الذى قد يظهر تفوقا فى مواد الدراسة وهنا يتحول العلاج إلى استغلال نواحى تعليمية . واختبارات التحصيل مفيدة أيضاً فى تقدير وتشخيص التدهور العقلى الناشىء عن أسباب عضوية أو انفعالية .

أنواع الاختبارات التحصيلية

بعض الاختبارات تقيس التحصيل فى مادة واحدة كالتاريخ أو العلوم . وهناك بطاريات من اختبارات التحصيل أعدت للفرق الإعدادية هذه البطاريات تتكون من عدة اختبارات فرعية وكل اختبار يقيس إحدى مجالات التعلم وما دامت الاختبارات الفرعية مقننة معاً ، فإن الاختبار يخبرنا بمعرفة الطفل لمادة بالمقارنة بقدرته فى المجالات الأخرى . وأكثر مجالات القياس شيوعاً المستوى الإعدادى هى القراءة والهجاية والقواعد والتاريخ ، والحساب . أما اختبارات الثانوى والمعاهد فقد أعدت لإعطاء

صورة عن النمو فى كل المجالات ، ولكن على وجه العموم ، زاد شيوع استخدام اختبارات المادة الواحدة عن استخدام بطاريات من اختبارات المواد .

وبالإضافة إلى اختبارات المسح التى تقدر التحصيل العام للشخص فى مجال واحد ، هناك اختبارات تشخيصية سيما فى القراءة والحساب . والاختبار التشخيصى لايحاول أن يحدد ، المستوى الذى وصل إليه الشخص ، بل يحدد نقاط القوة والضعف لدى التلميذ . فالاختبارات التشخيصية تركز على العملية التى يستجيب خلالها المفحوص ، أكثر مما تركز على نتيجة هذه العملية .

والعمليات التشخيصية أحياناً تستخدم فى تدريس أى موضوع . فالمدرس لا يستطيع أن يدرس بفعالية ما لم يعرف عيوب التلميذ . ولكن الاختبارات التشخيصية نادراً ما تكون مقننة لأنها تهتم أساسا بالمقارنة فى ذات الفرد وليس بمقارنة الفرد بالآخرين

وأغلب الاختبارات المنشورة تقيس المعرفة ، وتهمل جوانب السلوك الأخرى . وقد حدث تطور حديث فى القياس التعليمى هو إنشاء مقاييس عديدة جديدة لأهداف مثل : القدرة على تطبيق المبادئ ، ومهارات الاستدلال ، والقدرة على تفسير البيانات . هذا فضلا عن أن بعض المدارس قد استخدمت اختبارات الاهتمامات واختبارات الاتجاهات لدراسة تقدم التلميذ .

الفصل السادس عشر

اختبارات تحصيلية مقننة

عامـــة

- قياس التفكير .
- اختبارات النمو التعليمى .
- قياس مهارات المذاكرة .
- بطاريات للتحصيل العام .
- خاتمة .

الفصل السادس عشر

اختبارات تحصيلية مقننة

عامـــة

قياس التفكير :

برغم أن تدريب التفكير من أهم الأهداف العامة للتربية ، إلا أن الدراسات التربوية لم تعن به كثيراً . فقل عدد الاختبارات التحصيلية التى تقيس مدى التقدم فى عمليات التفكير نتيجة الاشتراك فى برنامج دراسى ، ومهارات التفكير بعد أن دربتها مختلف المؤسسات التربوية . وأكثر من هذا أن تعريف التفكير Thinking كان يتسع أحياناً فيشمل الكثير من المهارات والعمليات ، ويقتصر فى أحيان أخرى على مفهوم كالذكاء ، أو العامل العام ، أو النجاح الأكاديمى .

وقد قامت إحدى الدراسات ، التى استغرقت ثمان سنوات ومن ثم أصبح اسمها دراسة الثمان سنوات Eight Years Study ، بمحاولة لتصنيف عمليات التفكير إلى فئات قليلة محددة جيداً ومن ثم إلى إنشاء اختبارات موضوعية لكل فئة منها . وإن كان التصنيف غير شامل ، إلا أنه تضمن مظاهر من التفكير يمكن وصفها بالعبارات التالية :

١ ـ القدرة على تفسير البيانات .

٢ ـ القدرة على تطبيق المبادئ العلمية .

٣ ـ القدرة على تطبيق المبادئ المنطقية .

٤ ـ فهم طبيعة البرهان .

وكل فئة من الفئات الأربع تشمل منطقة واسعة من السلوك ، الأمر الذى دفع إلى تحديدها فى جوانب معرفة تبرز فى كل فئة . ففى حالة القدرة على تفسير البيانات ، قيست عناصر القدرات التالية :

(أ) القدرة على إدراك العلاقات بين البيانات :

١ ــ القدرة على عقد المقارنات :

٢ ــ القدرة على استخلاص عناصر مشتركة بين البيانات :

٣ ــ القدرة على قراءة البيانات .

٤ ــ القدرة على القيام ببعض العمليات الحسابية البسيطة بين هذه البيانات .

٥ ــ القدرة على فهم الرموز التى استخدمت وطرق عرض البيانات .

(ب) القدرة على معرفة حدود البيانات :

١ ــ القدرة على تحديد بيانات إضافية نحتاجها فى اشتقاق نتيجة معينة .

٢ ــ القدرة على التوقف عن إصدار الحكم حتى تتوفر كل البيانات اللازمة .

٣ ــ القدرة على الخروج بتعميمات من البيانات تكون قابلة للتحقيق Justifiable .

٤ ــ القدرة على تعيين التفسيرات التى لا يمكن تحقيقها .

٥ ــ القدرة على الخروج باستنتاجات جيدة تتعلق بالاتجاهات الواضحة بين البيانات .

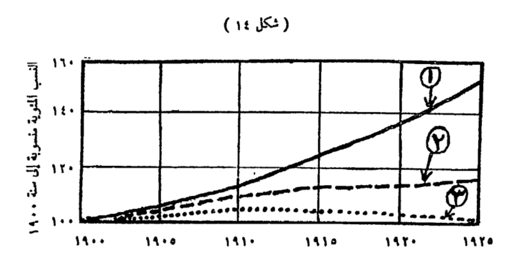
ولم تكن الفئات الفرعية السابقة كل فئات القدرة على تفسير البيانات ، فقد شملت مثلا القدرة على تقدير البيانات التى لا يمكن الاعتماد عليها ، أو القدرة على التقييم الناقد للاستنتاجات التى يقوم بها الآخرون ، ولكن

مثل هذه القدرات لم تُختَر للدراسة ، لأنه ليس عمليا أن يُدرس كل جانب من القدرة على تفسير البيانات :

أما الطرق التى صممت لقياس هذه القدرة فقد حوت بيانات عرضت بطرق مختلفة — بالرسم البيانى ، والجداول ، والألفاظ ، . . الخ . اختيرت هذه البيانات من مواد مقررة ، تشمل مجالات متعددة من الحياة ، وتمثل العديد من العلاقات بينها .

والشكل (١٤) يمثل طريقة عرض بعض البيانات التى تقيس القدرة على تفسير البيانات : والشكل البيانى متبوع بسلسلة من الجمل التى يطلب من المفحوص أن يحكم على صحتها مستعينا بالبيانات المعروضة بيانيا . والحكم

(شكل ١٤)



نموذج لأحد الأسئلة التى تقيس المهارة فى تفسير البيانات

الخط (١) مقدار الإنتاج الزراعى

الخط (٢) عدد السكان الزراعيين الذين فى سن العمل (الاستخدام)

الخط (٣) عدد العمال الزراعيين الذين يعملون (المستخدمين)

يكون على أساس مدى كفاية البيانات المعروضة وحدها فى تمكيننا من إصدار أحد الأحكام التالية :

(أ) هذه الجملة صحيحة .

(ب) يحتمل أن تكون صحيحة .

(ح) هذه الجملة غير صحيحة .

(د) يحتمل أن تكون غير صحيحة .

(ه) ليست هناك بيانات تمكننا من أن نصنف الجملة فى فئة من الفئات الأربع السابقة .

المشكلة الأولى : هذا الرسم البيانى يمثل الإنتاج ، والسكان ، وعدد المستخدمين فى مزارع الولايات المتحدة فى كل خمس سنوات ابتداء من سنة ١٩٠٠ حتى سنة ١٩٢٥ :

الجمل :

١ – النسبة بين الإنتاج الزراعى وبين العمال الزراعين تتزايد كل خمس سنوات بين سنة ١٩٠٠ وسنة ١٩٢٥ .

٢ – زيادة الإنتاج الزراعى بين سنة ١٩١٠ وسنة ١٩٢٥ ترجع إلى اتساع استخدام الآلات الزراعية .

٣ – متوسط عدد العمال الزراعيين المستخدمين خلال المدة بين سنة ١٩٢٠ وسنة ١٩٢٥ أعلى من عددهم خلال المدة بين سنة ١٩١٥ وسنة ١٩٢٠ .

٤ – يجب أن تنقذ الحكومة العمال الزراعيين العاطلين .

٥ – بين سنة ١٩٠٠ وسنة ١٩٢٥ ، زادت كمية الفواكه المنتجة فى مزارع الولايات المتحدة بمقدار ٥٠٪ .

٦ – خلال المدة الكلية بين سنة ١٩٠٥ وسنة ١٩٢٥ كان عدد

السكان الزراعين من أعمار يمكن تشغيلها أكبر من عدد الأفراد المطلوبين لتشغيل المزارع .

٧ — انخفضت الأجور المدفوعة للعمال الزراعين سنة ١٩٢٥ لأن عدد العمال الموجودين كان أكثر من العدد الممكن تشغيله :

٨ — عدد العمال الذين استخدموا سنة ١٩٢٥ زاد عن عدد المستخدمين سنة ١٩٠٠ :

٩ — منذ سنة ١٩٠٠ ، كانت الزيادة فى إنتاج كل عامل صناعى مساوية لزيادة إنتاج كل عامل زراعى .

١٠ — بين سنة ١٩٠٠ وسنة ١٩٢٥ ، زاد الإنتاج الزراعى بأكثر من ٥٠٪ .

١١ — زاد العمال الزاعيون إنتاجهم بعد سنة ١٩١٠ حتى يستفيدوا من الارتفاع السريع فى الأثمان .

١٢ — متوسط كمية الإنتاج الزراعى فى الفترة بين سنة ١٩٢٦ وسنة ١٩٣٠ كان أكثر منه فى الفترة بين سنة ١٩٢٠ وسنة ١٩٢٥ :

١٣ — بين سنة ١٩٠٠ وسنة ١٩٢٥ ، كانت هناك زيادة فى عدد السكان الزراعين من أعمار يمكن تشغيلهم فى المنطقة المتوسطة الغربية وهى أكبر المناطق الزراعية فى الولايات المتحدة :

١٤ — كان عدد السكان الزراعين من أعمار يمكن تشغيلها سنة ١٩٣٠ أقل من عددهم سنة ١٩٠٠ .

١٥ — إنتاج القمح ، وهو أكثر المحاصيل الزراعية فى الولايات المتحدة ، كان فى سنة ١٩١٥ مساوياً له فى سنة ١٩٢٥ :

وطريقة تصحيح هذا الاختبار معقدة بعض الشىء . وقد شملت

صحيفة التسجيل المستخدمة فى دراسة السنوات الثمانية أماكن لعدة درجات : وهذه الدرجات تتبع الفئات الأربع التالية :

١ — الدقة العامة : هذه الدرجة الكلية over-all تقيس مدى دقة استجابات التلميذ فى إتفاقها مع مفتاح تصحيح الاختبار . وقد اشتق المفتاح من أحكام مجموعة من الخبراء لكى يبين مدى صحة أو احتمال صحة الحكم فى كل حالة . ودرجة الدقة العامة يمكن أن تنقسم بدورها إلى الفئات الخمس التى يختار إحداها المفحوص حكماً يصدره على مدى كفاية البيانات فى الرسم فى دلالتها على الجملة المعطاة .

٢ — الحرص : هذه الدرجة تقيس ميل التلميذ إلى شدة الحرص فى تفسير البيانات . فإذا صنف جملة فى فئة « البيانات غير كافية » وكان الواجب أن تصنف فى فئة « تحتمل الصحة » ، كان المفحوص شديد الحرص . ودرجة الحرص Caution هى عدد الاستجابات التى بالغ فيها التلميذ فى الحرص على الاختبار .

٣ — ما وراء البيانات : هذه الدرجة تقيس الميل إلى زيادة التعميم من البيانات . أى أنها كانت عكس زيادة الحرص ، وبالتالى تقابل درجة ما وراء البيانات Beyond Data درجة الحرص . وعلى أى حال فإن التلميذ يستطيع أن يحصل على درجات فى المقياسين إذا عرف أن يتوسط بين الحرص الزائد وقلة الحرص ما دام التلميذ يحصل على درجة هنا ودرجة هناك .

٤ — الخطأ الفاحش : وهو عدد الأخطاء الواضحة ، وهى الأخطاء التى تحدث عندما يحكم المفحوص على الجملة بأنها غير صحيحة فى حين تكون صحيحة أو العكس .

وبالإضافة إلى الفئات السابقة ، تعطى درجات عن الإجابات المتروكة .

وهدف عملية التصحيح بمختلف مراحلها هو عرض الطريقة المميزة التي يفكر بها التلميذ عادة . فليس الاختبار إذن مقياسا ، له معيار التحصيل الكلي ، يقيم التلميذ في هذا الميدان . ويبين الجدول (١٢) أنواع التفسير التي يمكن أن تعطى للصفحات النفسية في هذا الاختبار . وهو يعرض درجات اختبار واحد والتفسير الذي يقترحه واضعو الاختبار لهذه الصفحة . وهم سميث Smith وتيلر Tyler والفريق الذي ساعدهما .

الفئـــــــــــة	درجات التلميذ س	متوسط الفصل
الدقة العامة	٥٤	٥٧
الدقه في الحكم بيحتمل الصحة ويحتمل عدم الصحة	٣٥	٣٨
الدقة في الحكم بعدم كفاية البيانات	٥١	٥٨
الدقة في الحكم بصحيح وغير صحيح	٧٦	٧٣
الحرص الزائد	٤٨	٢١
التعميم الزائد (ما وراء البيانات)	٤٣	٣٦
الأخطاء الفاحشة	١١	٨

(جـــدول ١٢)

يبين صفحة نفسية لتفسير بيانات اختبار دراسة السنوات الثماني

ولكي نفسر هذه البيانات الكمية يلزمنا تفكير محدد لمعرفة مدى قرب التلميذ س من متوسط فصله . فدرجاته أقل من المتوسط بدرجة صغيرة . وهذا يدل على نقص القدرة على التعرف على حدود البيانات . وبفحص أنواع الخطأ نتبين زيادة كبيرة عن المتوسط في الميل إلى الخروج عن

البيانات ، أو تقبل تعميات متجاهلا حدود البيانات . وليست هذه الدرجة من أعلى الدرجات فى الفصل (ولها دلالة ، ما دامت أغلب درجات التلميذ قريبة من متوسط الفصل) ، ولكنها أيضاً تدل على أن نسبة الأخطاء فى زيادة التعميم مقارنة بالأخطاء فى زيادة الحرص تزيد أيضاً عن متوسط الفصل (النسبة بين زيادة التعميم إلى زيادة الحرص لدى التلميذ = ٤٣ : ٤٨ ، ونفس النسبة بين متوسطى الفصل = ٣٦ : ٢١) ودرجة الأخطاء الفاحشة من أعلى الدرجات فى الفصل .

وأحد صعوبات مثل هذا النوع من الاختبارات هو أنه يحاول أن يقيس الكثير جداً فى وقت قصير جداً . والنتيجة هى أن الدرجات الجزئية ، لأنها تستند إلى فقرات قليلة ، لا يمكن الاعتماد عليها بدرجة كافية فى توجيه أو قياس التلميذ الفرد .

وقد أعد الفريق الذى قام بدراسة السنوات الثمانية عدة اختبارات أخرى وأشهرها الاختبار الذى يحاول قياس القدرة على تطبيق المبادئ العلمية . وهو اختبار شامل جيد إلا أن نظام تصحيحه أكثر تعقيداً من نظام تصحيح اختبار القدرة على تفسير البيانات . إذ تحسب للتلميذ ٢٠ درجة وبدراسة هذه الدرجات يمكن تحديد نوع الخطأ الذى يقع فيه التلميذ والعادات الفكرية غير المرغوب فيها .

ومن الجدير بالذكر أن هذه الدراسة التى بدأت سنة ١٩٣٢ اقترحت إعداد بعض الاختبارات ولكنها لم تستخدم إلا فى سنة ١٩٥٢ . وتمكينا للدراسات التى تستخدم هذه الاختبارات يسمح الناشر بإعطاء أصولها للقائمين بالدراسة ويطبعونها ولكن ليس من حقهم بيعها كما لا يقوم الناشر نفسه ببيعها . وهذا يدلنا على مدى حرص العلماء فى قصر استخدام الأدوات التى تحت التجربة على القائمين بالبحوث لا أن تباع فتفسد وتتحول إلى تجارة .

اختبارات النمو التعليمى

« اختبارات أيوا للنمو التعليمى » Iwa Tests of Educational Development يمكن أن تعد امتداداً « لاختبارات أيوا لكل تلميذ فى المهارات الأساسية » Iwa Every pupil Tests of Basic Skills الذى يقيس الفرق الدراسية ابتداءاً من الفرقة المقابلة للسنة الثالثة الابتدائية حتى الفرقة المقابلة للسنة الثانية الإعدادية . وإن كانت هـذه السلسلة من الاختبارات تشمل اختبارين للمعلومات العامة وليست لمهارات التفكير . والاختبارات حسب ما تدلنا أسماؤها تقيس النمو التعليمى العام ، وإن كان الاعتماد على الأسماء هو من قبيل الصدق الظاهرى إلا أنه يعطينا فكرة عما تقيس ، نحو الأهداف المرغوبة من التربية . إلا أن معنى هذه الأهداف هنا غير واضح تماماً .

وفى كراسة تعليمات الاختبار نجد مواصفات الاختبارات التى يشملها بالطريقة التالية :

الاختبار الأول : فهم المفاهيم الاجتماعية الأساسية — وهو اختبار يقيس معرفة التلميذ العامة وفهمه العام لموقف المجتمع الحديث . وكل تلميذ يحتاج إلى هذه المعرفة حتى يتمكن من التفكير الجيد فى المشاكل الحالية .

الاختبار الثانى : صحة الكتابة — وهو اختبار يحدد تمكن التلميذ من العناصر الأربع الأساسية فى الكتابة الصحيحة — التنقيط ، الاستخدام ، العنونة ، والهجاية . وهذا الاختبار لا يقيس القدرة على حفظ القواعد ، بل القدرة على استخدامها .

الاختبار الثالث : المعلومات العامة فى العلوم الطبيعية — وهو اختبار يحدد معرفة التلميذ العامة وفهمه للظواهر العامة ، ولأهم إسهامات المدنية الحديثة .

وهو كالاختبار الأول ، لا يتحدد بمقرر السنوات الدراسية فى المدرسة الثانوية فحسب . فقد اختيرت الفقرات على أساس ما يجب أن يعرفه الشخص الذى نال تعليما عاما .

الاختبار الرابع : القدرة على التفكير الكيفى – يقيس الاختبار القدرة على حل المشاكل العامة العملية . وهذه المشاكل تمثل فهم التأمين على الحياة ، ونظام التقسيط فى الشراء ، والقدرة على تقدير تكاليف بيت صغير وإصلاحه ومشاكل مشابهة .

الاختبار الخامس : القدرة على تفسير البيانات المقروءة فى الدراسات الاجتماعية – يقيس الاختبار قدرة التلميذ على تفسير وتقييم قراءات مختارة من الكتب المقررة ، والمجلات ، والجرائد . والقدرة على تقييم ، وفهم المتضمنات ، واستدلال المضمون الرئيسى للقطعة المختارة كلها جوانب يقيسها الاختبار .

الاختبار السادس : القدرة على تفسير المواد المقروءة فى العلوم الطبيعية – وهو كالاختبار الخامس ، يقيس قدرة التلميذ على تفسير وتقييم مختارات مقروءة . وهذا الاختبار يختلف عن الاختبار الخامس فقط فى المادة الدراسية موضوع الاختبار .

الاختبار السابع : القدرة على تفسير مواد أدبية – وهذا الاختبار يعرض مقتطفات مختارة من الأدب العالمى – فكل الأنواع الكبرى من الأدب ممثلة ، ويقيس التعرف على جو الكاتب وهدفه من بين ما يقيس .

الاختبار الثامن : المفردات العامة – وهو اختبار يقدم مقياسا لا لقدرة التلميذ على معالجة الكلمات فقط ، بل وأيضا لاستعداده العام للتعلم .

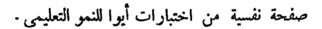
الاختبار التاسع : استخدام مصادر للعلومات — وهو اختبار يقيس قدرة التلميذ على الإفادة من مصادر المعلومات التى يغلب أن تشمل حل مشكلة معينة .

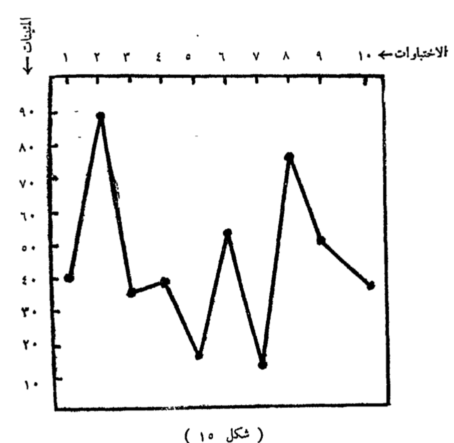
الاختبارات الخامس والسادس والسابع فى هذه السلسلة هى الاختبارات التى تعنينا أكثر من غيرها فى هذا المقام . ويجب أن نلاحظ أنها ليست اختبارات للمعلومات ، إذ كما تدل أسماؤها ، تؤكد على التفسير . والتفسير أحد جوانب قدرة التفكير التى تقاس بتقديم قطعة للتلميذ تتعلق بالمادة الدراسية موضوع القياس ويطلب منه أن يفسر معنى هذه القطعة . والأسئلة كلها مصاغة بطريقة الاختيار من عدة إجابات أو اختيار أفضل الإجابات من بين الإجابات المرفقة بالقطعة . Best Answer Form ولكن من الصعب أن نحدد بدقة ماذا تقيس مثل هذه الاختبارات . فهى قد تقيس المعرفة بموضوع المادة الدراسية ، وإن كان هذا الاحتمال صغيراً إذ أنها تطلب من المفحوص أن يستخدم أرضيته من المعلومات العامة تفسير القطع .

وبالرغم من أن هذه البطارية تؤكد على تفسير المادة فى مختلف ميادين الموضوعات الدراسية ، إلا أنها أيضاً تؤكد على قياس المعلومات الأساسية أو القاعدية إذ أن الأخيرة تعنى من يقيس النمو التعليمى . وليس هذا عيباً فى الاختبارات ما دام واضعوها يرون فى هذه الجوانب نقاطاً هامة يجب تقديرها .

وقد حــدث تطوير لاختبار « أيوا » كى يلائم أغراض الاختيار والتوجيه فى سلاح الطيران الأمريكى وعلى هذا أصبح اسمه « اختبار النمو التعليمى العام لمعهد سلاح الطيران بالولايات المتحدة » .

The United States Armed Forces Institute of General Educ-tional Development. (U.S.A.F.I.G.E.D.).





(شكل ١٥)

(١) اختبار فهم المفاهيم الاجتماعية الأساسية

(٢) « المعلومات العامة فى العلوم الطبيعية

(٣) « صحة الكتابة

(٤) « القدرة على التفكير الكيفى

(٥) « تفسير البيانات المقروءة فى الدراسات الاجتماعية

(٦) « تفسير المواد المقروءة فى العلوم الطبيعية

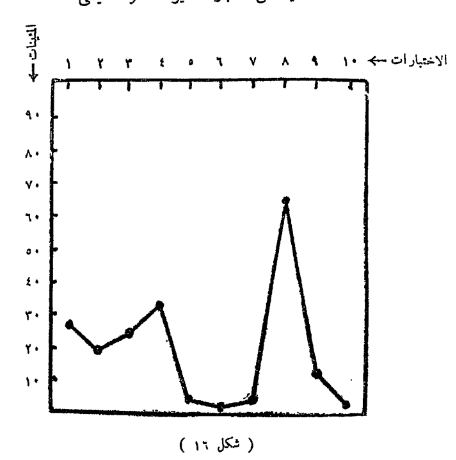
(٧) « القدرة على تفسير مواد أدبية

(٨) « المفردات العامة

(٩) الدرجة المركبة « الاختبارات من ١ — ٨ »

(١٠) استخدام مصادر المعلومات

صفحة نفسية من اختبارات أيوا للنمو التعليمى



(شكل ١٦)

(١) اختبار فهم المفاهيم الاجتماعية الأساسية
(٢) « المعلومات العلمة فى العلوم الطبيعية
(٣) « صحة الكتابة
(٤) « القدرة على التفكير الكيفى
(٥) « تفسير البيانات المقروءة فى الدراسات الاجتماعية
(٦) « تفسير المواد المقروءة فى العلوم الطبيعية
(٧) « القدرة على تفسير مواد أدبية
(٨) « المفردات العامة
(٩) الدرجة المركبة « الاختبارات من ١ — ٨ »
(١٠) استخدام مصادر المعلومات

وهو يستخدم لاختيار المتقدمين للإلتحاق بكلية الطيران من الحاصلين على دبلومات من المدارس الثانوية العسكرية ويتكون من أربعة اختبارات يقيس الأول صحة وسلامة التعبير ، ويقيس الثانى تفسير المواد المقروءة فى الدراسات الاجتماعية ، ويقيس الثالث تفسير المواد المقروءة فى العلوم الطبيعية ، ويقيس الرابع تفسير المواد الأدبية المقروءة . ولم تذكر فيه معاملات ثباته ، أما عن بحوث الصدق عليه فقد استخدم محك النجاح فى الكلية وكان عدد عينة الصدق ١٣٥ طالبا ومعامل الصدق ٧٢ر٠ ويصحح الاختبار بواسطة ماكينة التصحيح العالمية IBM والمعايير قومية . وعينة التقنين تمثل ستة مناطق جغرافية .

وهكذا نرى أن الاتجاه نحو قياس مهارات التفكير فى التربية أصبح مدعما بفضل دراسات واسعة النطاق مثل دراسة السنوات الثمانية واستخدام مقاييس النمو التعليمى فى عينات كبيرة كما حدث فى سلاح الطيران الأمريكى . والافتراض وراءها هو أن تعلم طرق التفكير الصحيحة الفعالة يمكن الفرد من التكيف والتوافق فى أكبر مدى من المواقف .

على أن المدرس الذى يستخدم أداة من هذه الأدوات يجب أن يحلل ما يبدو أن الاختبار يقيسه فيقرأ المدرس الاختبار ، ثم يجيب عن أسئلته ، ثم يسائل نفسه قائلا :

١ — ما عدد فقرات الاختبار التى تقيس المعرفة بمفردات موضوع القياس ؟

٢ — ما عدد الفقرات التى تقيس القدرة على اشتقاق الأدلة من البيانات ؟

٣ — ما هو عدد الفقرات التى تسأل التلميذ أن يخرج بتعميمات من البيانات ؟

فإذا ساءل المدرس نفسه بهذه الطريقة كان فى موقف يمكنه من الحكم على صلاحية الأداة فى قياس الأغراض التى يضعها فى ذهنه ؟

قياس مهارات المذاكرة

يؤكد المدرسون فى كل مستويات التعليم أهمية تحسين مهارات المذاكرة ، ويعتبرها البعض ضمن جوانب مهارات التفكير . وتكثر الشكوى ، فى أيامنا هذه ، من عدم كفاية الطرق التى يتبعها التلاميذ فى المذاكرة ، وخاصة فى مراحل التعليم الأولى . ولا يهمنا أن نلقى تبعة هذا على عواهل المدرس أو التلميذ قدر ما يهمنا أن تقوم أبحاث منظمة علمية لتحسين مهارات المذاكرة يقوم بها المدرسون والإخصائيون داخل وخارج الفصل . وسنتحدث عن مهارات المذاكرة بتقسيمها إلى الفئات الرئيسية التالية :

١ ـ مهارات القراءة : وهذه المهارات متضمنة غالباً فى المهارات اللغوية ، ولكن المتخصصين فى مهارات المذاكرة يهتمون عند التعرض لمهارات القراءة بتشخيص أسباب صعوبات القراءة .

٢ ـ مهارات الكتابة : بما فى ذلك إعداد التقارير . وهى تتضمن مهارات التنقيط والعنونة وتنظيم عرض المفاهيم المتضمنة .

٣ ـ مهارات الإصغاء : وتتضمن مهارات الإصغاء Listening إلى جوار حسن الاستماع والانتباه مهارات كتابة المذكرات والمحاضرات والملخصات .

٤ ـ مهارات تنظيم الأفكار : وهى متضمنة فى مهارات اللغة . ولكنها تبرز أثناء القراءة .

٥ ـ مهارات التخطيط : ونعنى بها عمل الجداول وتخطيط عمليات المذاكرة :

٦ - مهارات إجابة الامتحانات : وهى تتضمن مهارات الكتابة ومهارات استخدام الأدلة والنقط الهامة فى الإجابة عن الأسئلة الموضوعية :

٧ - مهارات جمع البيانات : ومهارات الإفادة من المكتبات .

وقد قام عدد من العلماء بتقديم قوائم يسجل فيها التلميذ نوع الصعوبات التى يواجهها فى المدرسة . وهذه القوائم تفيد فى القيام بالمقابلات الشخصية Interviews مع التلاميذ لتوجيههم وإرشادهم ، إذ أن التلميذ لا يكون عارفا ما هى أسباب مشاكله ومصادرها :

وأفضل قائمة من هذا النوع هى قائمة مراجعة المشاكل التى أعدها مونى Mooney : وهى تستخدم على نطاق واسع الآن خاصة فى المدارس الثانوية والإعدادية فى الخارج . وقد أعدت من هذه القائمة صور كثيرة أولاها للمدارس الإعدادية ، وأخرى للثانوية ، وثالثة لطالبات التمريض ، ورابعة للشباب من القرويين وإن كانت الصورتان الأخيرتان لا تستخدمان الآن . كما أعدت صورة للراشدين وأخرى للجامعيين .

تم إنشاء هذه القوائم بناءاً على الكثير من المواد ذات الدلالة التى جمعت عن مشاكل فعلية . قال تلاميذ المدارس الثانوية إنهم يواجهونها ، وقد تم جمع هذه المواد عن طريق المقابلات الشخصية مع التلاميذ ، وتحليل سجلات الحالات ، وتحليل فقرات كتبها التلاميذ عن مشاكلهم ، ومن التعبيرات الكتابية للتلاميذ عن حاجاتهم التربوية ، ومن نتائج بحوث سابقة مماثلة .

وتتضمن القائمة الكثير من المشاكل مثل الخجل - الخضوع - التمركز حول الذات - ضخامة المسئولية العائلية : ومجموع فقرات القائمة ٣٠٠ فقرة . وبعضها يعالج عدم كفاية مهارات المذاكرة ، والعادات الخاطئة

فى المذاكرة ، ومصادر الفشل الدراسى . ويطلب من الطالب أن يقرأ القائمة ببطء وأن يضع خطا تحت الفقرات التى تشير إلى مشاكل يعانى منها . وبعد أن يكمل المفحوص هذه الخطوة الأولى ، يراجع الفقرات ويضع خطوطا ودوائر حول أرقام بعضها الذى يعده « أكثر ما يتدخل فى حياته » . وأخيراً يسأل أن يكتب ملخصاً لمشاكله . ولا تصحح القائمة بل تعد كشفاً لمشاكل التلميذ .

وقد اقترح مونى عدة مناهج لبحث صدق قائمة مراجعة المشاكل ؛ والمناهج هى :

١ ـ يسأل الطالب الذى أجاب عن القائمة السؤال التالى :

« هل تشعر أن فقرات القائمة التى وضعت علامة عليها تعطى صورة كاملة عن مشاكلك ؟ » .

وقد وافق ٩٢٪ ممن أجابوا عن القائمة على أن هذه القائمة تعطى صورة كاملة تقريبية عن مشاكلهم واعتبروا ذلك دليلا على صدقها ؛ والصدق هنا يعنى أن إجابة المفحوص تعكس صـــورة عما يعتقد هو أنها مشكلات تعترضه . وهذا لا يعنى بالضرورة أن إجاباته ترتبط بالتشخيص الإكلينيكى ارتباطا كاملا . والسبب أن استبصار التلميذ بطبيعته الخاصة محدود . وإذا نمى تبصر التلميذ قد يرى فى إدراكه لمشكلاته سابقاً رأياً آخر فيكتشف مثلا أن ما كان يعانى منه فعلا هو مشكلات أخرى غير التى وافق عليها فى القائمة . وعلى هذا فالصدق بالمعنى الأول ليس الصدق الذى يعنينا كمشتغلين بالقياس . ولكنه صدق رجل الشارع ، أى صحة أو عدم كذب الفقرات واتفاقها مع إدراك التلميذ لنفسه .

٢ ـ كما اقترح مونى دراسة صـــدق قائمته بملاحظة مدى اتفاق الإجابات مع المشاكل الشخصية للمجيبين وكيف أن الإجابات تتغير بتغير المشاكل . ولبحث هذا الدليل قام مونى بالتجربة الآتية :

أتى بمجموعتين تكونت الأولى من ٧٠ طالباً والأخرى من ٤٦ طالبة وأعطيت لهم قائمة المراجعة . وبعد ٩ أيام طلب منهم أن يضعوا علامات على المشاكل التى حلوها وعلى المشاكل التى جدت عليهم وكانت هذه العلامات توضع فى إحدى الاستمارات المصغرة عن هذه القائمة . ثم أعيد إجراء القائمة وعندئذ ظهر أن ٨٣٪ من المشاكل التى وضعت عليها علامات فى الاستمارات المصغرة وضعوا عليها أيضاً علامات فى القائمة : ولكن هذا الإتفاق قد لا يعنى صدق القائمة بقدر ما يعنى درجة عالية من الإتساق بين الفرد ونفسه فى مرتى الإجابة عن أسـئلة تذهب إلى نفس المجرب .

٣ ــ أما الطريقة الثالثة فهى أفضل من سابقتيها إذ تغلبت على النقد الموجه إليهما من أن التلميذ لا يعرف بدقة مصادر مشكلاته فاتفاقه فى الإجابات أو اختلافه بعد فترة لا يعنى أن مشكلاته تغيرت ولكنه يعنى أن إدراكه لهذه المشاكل قد تغير . فإذا كان هذا الإدراك نفسه غير صادق تماماً فالأحرى أن تكون عمليات بحث صــدق القائمة المعتمدة عليه غير مؤكدة أيضاً .

والطريقة الثالثة الفضلى هى اختيار مجموعات ثبت من دليل آخر صادق أن لها مشكلات محددة ثم تطبيق قائمة المراجعة عليها لتبين ما إذا كانت الفقرات تعكس المشكلات أم لا . وقد قام بدراسة من هذا النوع ستودجيل Stodgill وبينتون Benton حيث طبقا قائمه المراجعة على مجموعتين تتكون كل منهما من ٣٥ طالباً جامعياً . وإحدى المجموعتين فصل ينال تلاميذه علاجاً فى مهارات المذاكرة والأخرى فصل ينال تلاميذه علاجا فى الصحة النفسية .

وقد زاد تكرار إجابات تلامذة المجموعة الأولى عن فقرات مثل « لا أعرف كيف أذاكر بطريقة جيدة » ، « أخشى الرسوب فى الكلية » ، « لا أعمل شيئاً بكفاءة » ، « المدرسون لا يهتمون بالتلاميذ » ، « احتاج

إلى أن أعرف القدرات المهنية » ، « غير قادر على التركيز » ، بطىء القراءة » ، . . الخ . أما تلامذة فصل الصحة النفسية فقد زاد تكرار إجابات تلاميذه عن فقرات مثل « أشعر بالاضطهاد » ، « مضطرب فى عقائدى الدينية » ؛ « إن الآباء يرهقون الأبناء بكثرة مطالبهم » ، « لا أجد وقتاً كافياً لنفسى » ، « أحتاج إلى برامج علاجية أخرى » :

بينت هذه الدراسة أن المشاكل التى وضع الطلبة تحتها خطوطا فى قائمة المراجعة هى المشاكل التى يعانون منها . ومع ذلك يجب أن نكون حذرين جداً عند تفسير القوائم الفردية فنحن نعلم أن المريض لا يكون متبصراً بسبب مشكلاته . فهو يربط القلق بهذا الموضوع أو ذاك ويتلمس فى الأسئلة أو الفقرات إشارة إلى هذا الموضوع . وعلى ذلك نجد أنه يعانى من القلق ويجيب عن فقرات القلق بالإيجاب أى الموافقة . ولكنه لا يتعرف على سبب هذا القلق ، حين أننا يجب أن نبحث دائماً عن مصادر هذه الأعراض أو المشاكل وليس عن الموضوعات التى ارتبطت بها : فالقائمة إذن تعين على تشخيص المشكلات وليس على تحديد أسبابها ومصادرها .

بطاريات التحصيل العام
فى المرحلتين الابتدائية والإعدادية

تستخدم بطاريات التحصيل العام general Achievement ابتداء من المستوى الابتدائى حتى الدراسات العليا . إلا أن أفضل استخدام لها فى المرحلة الإعدادية . وأغلب هذه البطاريات تعطى صفحات نفسية تمثل درجات مختلف المواد الدراسية كما أنها تعطى درجة كلية عن البطارية ككل . وهذه البطاريات تختلف فيما بينها فى عدة خصائص سواء من حيث المواد التى تقيسها أم المستويات المقاسة . كما تختلف فى معاييرها

ومعاملات صدقها وثباتها . ولذا يصعب أن نأخذ بطارية واحدة لنمثل بها خصائص هذا النوع من البطاريات .

وقبل أن نعرض هذه البطاريات ننبه إلى أنها لا تصلح للاستخدام فى بلادنا إلا بعد الكثير من التعديلات . وأولى بواضعى الاختبارات أن يصمموا بطاريات واختبارات فيها أصالة Originality تمثل دراسة شاملة لحاجاتنا وإمكانياتنا التربوية وتناسب المواد الدراسية فى مختلف مستويات التعليم العام والخاص . كما ننبه إلى أن أغلب ما سنعرض يستخدم فى أمريكا ونحن نقدمه مع بعض التعديلات حتى يمكن أن يفهم خلال أوضاعنا فيحفز البعض على استخدامه والآخر على تصميم اختبارات من هذا النوع .

أولا : اختبارات متروپوليتان التحصيلية : إن اختبارات متروبوليتان التحصيلية Metropolitan Achievement Tests من أشيع بطاريات التحصيل استخداما فى مستوى المرحلة الإعدادية وإن كانت تستخدم فى السنوات الدراسية ابتداء من الأولى الابتدائية حتى نهاية المرحلة الإعدادية . وهى تحتوى ٥ بطاريات ولكل بطارية ٤ أو ٥ صور متكافئة . وهى اختبارات قوة . ويحتاج تطبيق كل البطارية إلى حوالى ٤ ساعات . وكل الاختبارات التى تشملها البطارية الواحدة مطبوعة فى كتيب واحد .

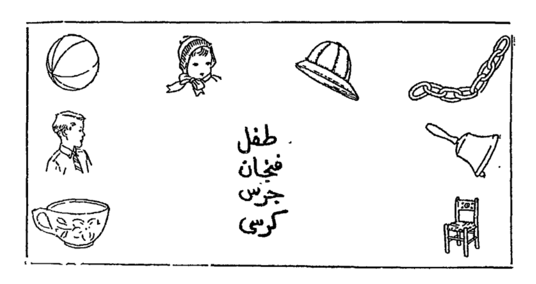
طبقت الاختبارات على نطاق واسع واختيرت عينه قومية وحسبت الأعمار المنوالية لكل فرقة دراسية . والمعايير التى استخدمت هى معايير العمر المنوالى والعمر العقلى والفرقة الدراسية والتوزيع المئينى للفرقة الدراسية . وقد حسبت معاملات ثبات كل بطارية على حدة بالطريقة النصفية وكانت معاملات الثبات تتراوح بين ٨ وإلى ٩٧ر٠٠ وكان المحك فى بحث الصدق هو صدق المنهج الدراسى وارتفاع درجات المفحوصين فى الاختبارات بتقدم فرقهم الدراسية . وعلى نفس الأساس اختيرت الفقرات إذ قورنت

النسب المئوية للناجحين بفرقهم الدراسية كما قورنت درجات الأفراد فى الفقرات واحدة بواحدة بدرجاتهم فى الاختبارات .

وكى يتضح مدى ومحتوى هذه السلسلة من البطاريات نعرض أول بطارية منها وهى تقيس السنة الأولى الابتدائية . وتتكون من الاختبارات التاليـــة :

الاختبار الأول : القراءة ــ صورة الكلمة Word Picture ويتكون من ١٤ مجموعة من الصور وكل مجموعة بها ٨ صور . ووسط كل مجموعة منها عدة كلمات أو عبارات مكتوبة يطلب من المفحوصين أن يضعوا علامات توصل كل كلمة أو عبارة مكتوبة بالصورة التى تناسبها . ويقوم الفاحص بمساعدة المفحوصين فى نطق الكلمات المطبوعة .

الاختبار الثانى : القراءة ــ التعرف على الكلمة Word Recognition ويتكون من ٢٦ مجموعة من الكلمات كل مجموعة من ٤ كلمات تتشابه فى النطق و الهجاية . وينطق الفاحص بكلمة منها فيضع المفحوصون



(شكل ١٧)
فقرة من اختبار صورة الكلمة فى بطارية متروبوليتان التحصيلية

علامة على هـــذه الكلمة فمثلا الكلمات فى الصف هى Pig big bid dig وينطق الفاحص كلمة big ويترك للتلاميذ فرصة ١٥ ثانية لوضع علامة على هذه الكلمة ، هكذا فى كل صف .

الاختبار الثالث : القراءة — معنى الكلمة Word meaning ويتكون الاختبار من ١٢ مجموعة من الكلمات كل فى صف وتتكون المجموعة من ٦ كلمات تشترك ثلاثة منها فى صفة . ويقول الفاحص هذه الصفة ويطلب من التلاميذ وضع علامات على الكلمات التى تتوفر فيها هذه الصفة فمثلا نجد الصف الأول : منزل — ولد — حجر — فأر — رجل — أب ، ويطلب الفاحص أن يضع التلميذ علامة على أى كلمة فى هذا الصف تدل على شىء متكلم ويساعد التلاميذ فى المحاولتين الأولى والثانية بقوله مثلا « صحيح ؛ ولد — رجل — وأب » .

الاختبار الرابع : الأرقام numbers — ويتكون من ٣٦ صفا من الصور أو الأرقام ويطلب من المفحوص أن يضع علامة حسب التعليمات . مثلا صف به ٩ رجال أمام باب ويطلب من المفحوص أن يضع علامة على الشخص الخامس . أو عدة أرقام ويطلب من المفحوص وضع علامة على الرقم ٦ مثلا . ثم ثلاثة صفوف لعمليات الجمع البسيطة عددها ١٥ عملية مثل ($\frac{١}{٦}$) وصف لعمليات طرح عددها ٥ مثل ($\frac{٣}{-٢}$) .

وهذه الاختبارات تتطلب القراءة . وهى شبيهة اختبارات الذكاء العام



(شكل ١٨)
نقرة من اختبار الأعداد فى بطارية متروبوليتان التحصيلية

فى المرحلة الابتدائية وخاصة فى اختبارات الأعداد وإن كانت أصعب فى البطارية منها فى اختبارات الذكاء إذ أنها فى الأولى تطلب من التلميذ قراءة وطرح وجمع كسور عشرية . والبطارية تقيس حتى نهاية المرحلة الإعدادية وهنا نجد اختبارات فى القراءة والمفردات وأسس الحساب ومسائله واللغة وأدبها والمواد الاجتماعية (مثل التاريخ والجغرافية) والعلوم والهجاية .

وفى اختبارات القراءة ، فى نهاية المرحلة الإعدادية ، يقرأ التلميذ قطعة صغيرة ثم يجيب عن أسئلة تتعلق بمضمونها ومعناها . هذا الاختبار واختبار المفردات قريبان من اختبارات الذكاء . أما بقية الاختبارات فتختلف قليلا وتقترب كثيراً من محتوى المنهج الدراسى .

ثانياً اختبار ستانفورد للتحصيل : Stanford Achievement .

يعتبر اختبارات ستانفورد للتحصيل من أقدم البطاريات التعليمية المقننة . وقد نشر لأول مرة سنة ١٩٢٣ وروجع سنوات ١٩٢٩ ، ١٩٤٠ ، ١٩٥٣ وفى كل مراجعة حدثت تعديلات كثيرة . والطبعة الحالية (نقصد طبعة سنة ١٩٥٣ التى تصبح فعالة حتى ١ يناير سنة ١٩٥٤) تتكون من ٤ بطاريات وكل بطارية لها ٥ صور متكافئة . وتقيس البطاريات الأربعة الفرق الدراسية من الثانية الابتدائية حتى نهاية المرحلة الإعدادية . وفى المراجعة التى نتكلم عنها ، أصبح محتوى الاختبارات مسايرا لأحدث تعديلات المناهج الدراسية كما سهل التصحيح وبات موضوعيا .

وقد اشتقت معايير جديدة بعد تطبيق البطاريات على ٤٥٠٠٠٠ تلميذاً وتلميذة يمثلون ٣٦٣ مجموعة . واستخدمت الدرجة ك K. Score وهى تشبه الدرجة المعيارية المستخدمة فى بطاريات المتروبوليتان . وبدلا من أن تشتق هذه الدرجات من المنحنى الاعتدالى اشتقت من منحنيات ملتوية

تمثل توزيع تلاميذ كل فرقة دراسية على حدة بالنسبة للنجاح المدرسى . والمعايير من نوع العمر المتوالى ، ومعيار الفرقة الدراسية ، والتوزيع المئينى لدرجات تلاميذ كل فرقة دراسية على حدة .

حسبت معاملات ثبات اختبارات هـذه البطارية بالطريقة النصفية وتراوحت من ٦٦ر٠ إلى ٩٦ر٠ وعند وجود معامل ثبات أقل من ٨٠ . وكانت الفقرات تراجع على أساس أن نسبة الناجحين فى الفقرة يجب أن تزيد بتقدم المفحوصين تعليميا .

والواقع أن هناك تشابه بين بطاريتى المتروپوليتان وستانفورد للتحصيل ينحصر أساسا فى :

١ ــ أنهما تقيسان نفس الفرق الدراسية تقريبا .

٢ ــ صعوبة الأسئلة متقاربة بين الاثنين .

٣ ــ التعليمات متشابهة إلى حد ما .

٤ ــ كل من البطاريتين به بطارية جزئية للمواد التى تتطلب مهارات مثل الحساب والقراءة .

٥ ــ نشرتهما شركة الكتاب العالمى World hook Co. .

إلا أننا نلمس أيضاً أوجه الخلاف الآتية :

١ ــ بطارية متروپوليتان أحدث ولكن مراجعتها أقل .

٢ ــ بطارية ستانفورد تمثل أحدث الاتجاهات التربوية على عكس بطارية المتروپوليتان .

٣ ــ بطارية المتروپوليتان بها اختبار منفصل للجغرافية وآخر للتاريخ بينما فى ستانفورد اختبار واحد للاثنين معاً .

٤ ــ اختبارات ستانفورد تمتاز بوجود اختبارات لدراسات مهارات التعلم والمذاكرة الحديثة مثل قراءة الخرائط والرسوم البيانية والجداول واستخدام المعاجم وغير ذلك من المصادر المرجعية .

فى المرحلة الثانوية

أولا : اختبارات كاليفورنيا التحصــيلية : California Acievement

تقيس من السنة الأولى الابتدائية حتى نهاية المرحلة الثانوية وتشمل ٤ بطاريات لكل منها ٢ أو ٣ صورة متكافئة . ويرى واضعوها أن المواد الدراسية مثل العلوم والجغرافية والتاريخ يختلف محتواها من مدرسة إلى أخرى ومن منطقة إلى أخرى بينما يعم تدريس المواد التى تطلب مهارات . وكل البطاريات تشترك فى قياس المجالات الخمسة الآتية :

١ — مفردات القراءة .

٢ — فهم المادة المقروءة .

٣ — الاستدلال الحسابى .

٤ — الأسس الحسابية .

٥ — اللغة .

وتفسر الدرجات عن هذه الاختبارات حسب معايير الفرقة الدراسية ، أو العمر ، أو التوزيع المئينى للدرجات داخل الفرقة الدراسية الواحدة . وقد جمعت البيانات المعيارية من عدد كبير من المناطق الجغرافية . وهى اختبارات تقيس القوة أكثر مما تقيس السرعة . وكل بطارية يمكن أن تعطى فى أكثر من جلسة واحدة .

ولهذه البطاريات ميزة استخدامها كأدوات تشخيصية . ولذا تحسب لكل تلميذ ٢٠ درجة فرعية فى أجزاء الاختبار الواحد ، فضلا عن ٥ درجات كل عن اختبار . وتمثل الدرجات فى صفحة نفسية . وبهذه الطريقة مثلا ينقسم اختبار مفردات القراءة إلى أقسام مثل : مفردات حسابية ، وعلمية ، ومفردات الدراسات الاجتماعية والأدبية . كما ينقسم اختبار اللغة إلى التنقيط ،

ووضع الفواصل ، واختيار الجمل ، والكلمات واختبار لأجزاء الحديث وآخر للهجاية وآخر للخط و . . الخ .

أما اختبار الاستدلال الحسابي فينقسم إلى مفاهيم الأرقام ، والرموز ، والقواعد ، والأرقام والمعادلات ، والمسائل : واختبار العمليات الحسابية ينقسم إلى عمليات للجمع والطرح ، والضرب ، والقسمة وقد ينقسم أكثر من هذا إلى جمع بسيط ومركب وإيجاد العوامل المشتركة ، . . . الخ .

ونلاحظ ، بناء على هذا العرض ، أن هذه الاختبارات تحاول أن تقيس الكثير جداً بالقليل جداً من الفقرات التي لا تغطي هذا العدد الضخم من القدرات . ولكن الاختبارات الخمس ثابتة ويعتمد عليها في الغرض من إنشائها في كل المستويات الدراسية حتى نهاية المرحلة الثانوية . وقد أوجدت معاملات ثبات كل اختبار من الخمسة ولم توجد معاملات الاختبارات الفرعية إذ أن بعضها قصير جداً .

ثانياً : بطارية المقرر الأساسي للمدرسة الثانوية : Essential High School Content Battery.

ثالثا : الاختبارات التعاونية للتحصيل : The Cooperative General Achievement Tests (C.G.A.T.)

رابعا : بطاريات النمو التعليمي : وقد سبقت الإشارة إليها .

خاتمة :

ومن عرضنا السابق لكل بطاريات واختبارات التحصيل العام والنمو التعليمي يتضح لنا :

١ – أن الكثير من هذه الاختبارات لا يقدم بعض معالم القياس الهامة فلا نجد مثلا معاملات الصـــدق التجريبي لاختبارات أيوا واختبار U.S.A.F.I.G.E.D. ؛ كما لا نجد في الاختبارين أيضاً معاملات الارتباط

الداخلية بين الاختبارات الفرعية . ومعاملات ارتباط هذه الاختبارات ؛ مع اختبارات تحصيلية أخرى مقننة .

٢ ــ أن هذه الاختبارات ، عموماً ، تقيس وظيفة قريبة مما تقيسه اختبارات الاستعدادات المدرسية . وكلا النوعين يأخذ محكا له من النجاح الأكاديمى .

٣ ــ أن بعض الاختبارات التحصيلية تعنى بقياس المهارات المعينة التى تدخل فى نطاق الاستعدادات الخاصة والاستعدادات الفارقية .

وباختصار فإن الاتجاه الحديث فى إنشاء وتطوير الاختبارات التحصيلية قد أفاد كثيراً من تقدم النظريات التربوية . كما أفادت منه المناهج الدراسية .

إلا أنه ، كما أسلفنا ، لا تتوحد الخبرات التعليمية بين التلاميذ غالباً نتيجة لتغيرات النظم التعليمية ولاختلاف امتحانات المناطق وتفاوت النجاح الأكاديمى بين التلاميذ أفراداً . ولهذا فعند تطبيق اختبارات النمو التعليمى العام على أفراد مختلفين فى خبراتهم التربوية يجب أن يركز الاهتمام على وظيفتين نقيم بهما التلاميذ موضوع القياس هنا وهما :

١ ــ القدرة على إدراك وفهم المعلومات :

٢ ــ مدى توحد خبراتهم وأثر تعليمهم السابق :

الفصل السابع عشر

اختبارات تحصيلية مقننة

خاصـــة

- القراءة :
- الرياضيات .
- العـــلوم .
- الجغرافية .
- مجالات أخرى .
- مقاييس الإنتاج .
- أغراض تعليمية أوسع .
- الاختبارات التحصيلية المهنية .

الفصل السابع عشر

اختبارات تحصيلية مقننة

خاصة

القراءة :

تستحق دراسة القراءة اهتماماً خاصاً لعدة أسباب : أولها أن القراءة هي الأداة الرئيسية التي عن طريقها يتعلم الفرد كل المواد : وثانيها أن عدد اختبارات القراءة التي أعدت يزيد كثيراً من حيث العدد ومن حيث التنوع أيضاً عن غيرها من الاختبارات . وثالثها أنها أداة هامة شائعة الاستخدام في الإرشاد والتوجيه النفسي وفي ميادين الخدمات الإكلينيكية المختلفة . هذا فضلا عن أن اختبارات القراءة تعرض لنا الكثير من المشاكل التي تصادفنا في عمليات إنشاء اختبارات التحصيل المقننة الخاصة .

تعريف القراءة :

قد يبدو لنا من أول وهلة أن القراءة مهارة بسيطة واضحة التعريف ميسورة القياس . ولكننا إذا ما استعرضنا اختبارات القراءة المشهورة وجدناها تختلف فيما تقيس من جوانب السلوك اختلافا يزيد عما نجده في قياس أي مهارة أخرى . ولم يختلف مصممو اختبارات القراءة على جوانب السلوك التي يجب أن تتضمن خلال قياس القراءة فحسب ، بل وإنه أيضا اختلفوا في تعريفاتها . اختلفوا في تحديد مدى تأثير عوامل السرعة ، والفهم ، والمفردات . . . إلخ : كما اختلفوا في تحديد أهداف قياسها .

وقد قام تراكسلر Traxler بفحص ٢٤ اختباراً للقراءة ووجد أنها تقيس ٤٨ مهارة مختلفة .. ولا يعني هـذا أن القراءة قدرة أو مهارة

أو استعداد مركب من عدد ضخم من العوامل النوعيـة أو الخاصة أو الفرعية . إذ ادعى أحد الاختبارات أنه يقيس عدداً من مهارات القراءة المختلفة تماما ، ولكن فحص معاملات الارتباط بين درجات المفحوصين فى هذه المهارات بين أنها فى الواقع وظيفة واحدة تقاس تحت أسماء متعددة .

وأحد طرق توضيح ما إذا كانت قدرات القراءة مختلفة حقا هو التحليل العاملى . فقد وجدت العوامل الآتية شائعة فى ٢٥ اختبارا للقراءة ومهارات المذاكرة : الميل إلى القـــراءة بعناية (وهو اتجاه أو جهاز استجابات) ، والاستدلال الاستقرائى ، وسرعة القراءة ، والقدرة اللفظية ، والمفردات ، وقراءة الخرائط . ولهذا يجب عند الحاجة إلى استخدام اختبار للقراءة أن نعرف مضمون الاختبار ، وأن نحدد أى قدرة قرائية نريد أن نقيس ، وأن نرجع إلى مراجع يعتمد عليها فى تحليل الاختبارات التى سنستخدمها .

اختبارات القراءة إذن من أهم الاختبارات التحصيلية التعليمية . وهى تنقسم ، كما تنقسم كل اختبارات التحصيل المقننة الخاصة ، إلى ثلاثة أقسام هى :—

١ ـ اختبارات التهيؤ readiness .

٢ ـ اختبارات المسح survey .

٣ ـ اختبارات التشخيص diagnostic Tests .

وتطبق اختبارات التأهب عند قبول التلاميذ بالمدارس ، لتحديد ما إذا كان الطفل ذا مهارات ومعلومات تمكنه من تعلم القراءة . أما اختبارات المسح فهى تستخدم المقاييس الكلية Overall measures فى معرفة مستوى القدرة القرائية . بينما تستخدم اختبارات التشخيص فى تعيين نقاط الضعف والقوة التى يعانى بسببها الفرد من مشكلات تعليمية أو نفسية . ولهذا نجد اختبارات تشخيص القراءة تهتم بالتحليل الكيفى لجوانب قدرات القراءة

ودراسة الأخطاء النوعية المعينة التي غالبا ما يقع فيها تلميذ بذاته ومدى تكرار الوقوع فى هـذه الأخطاء . ومن ثم يمكن بالاستقصاء التالى معرفة مصادر هذه العيوب وأسبابها بربطها غالبا بنوع من الخبرات المؤلمة أو بإرجاعها إلى عدة عوامل متفاعلة متكاملة كالأسرة والمدرسة و . . الخ .

« اختبارات التهيؤ فى القراءة »

برغم أن اختبارات التهيؤ تنتمى إلى اختبارات الاستعداد ، إلا أن استخدامها فى المواقف التعليمية وارتباطها بأنواع اختبارات القراءة الأخرى خاصة وبأنواع اختبارات التحصيل فى الميادين الأخرى عامة ، يحتم الحديث عنها هنا . هذا برغم أن اختبارات القراءة ، خاصة فى مستوى المدرسة الابتدائية ، والمدرسة الإعدادية ، تقترب من اختبارات الذكاء العام من حيث الوظيفة المقاسة ، ولكن اختبارات التهيؤ فى القراءة تهتم بقياس القدرات التى يعتقد أن لها أثر فى تعلم القراءة وفى النجاح المدرسى .

واختبارات التهيؤ فى القراءة تقيس الوظائف الآتية :

١ ــ التمييز البصرى والسمعى .

٢ ــ الضبط الحركى .

٣ ــ الاستدلال اللفظى :

٤ ــ المفردات .

٥ ــ المعلومات العامة .

ومن أهم اختبارات التهيؤ فى القراءة وأكثرها استخداما نذكر الاختبارات الآتية :

١ ــ اختبارات جيتس للتهيؤ فى القراءة .

Gates Reading Readiness Tests

٢ ــ اختبارات المتروپوليتان للتهيؤ فى القراءة :

Metropolitan Reading Readiness Tests

وهما اختباران جماعيان . أما الاختبارات المشهورة التى تستخدم فى المواقف الجماعية والفردية أيضاً فأهمها :

اختبارات مونرو للاستعداد للقراءة : The Monroe Reading A ptitude Tests

كما أن هناك اختبارات فردية للتهيؤ تهتم بالمقاييس النفسية والجسمية مثل بطارية بيتس للاستعداد للقراءة :

Betts Ready to Read Battery.

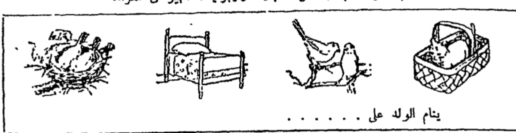
وتختلف هذه الاختبارات عن بعضها ، فضلا عما ذكرنا من أوجه ، فى المدى والوظائف المقاسة . كما تختلف أيضاً فى بعض الخصائص الإحصائية والقياسية الأخرى . فمثلا نجد أن معاملات صدقها تتراوح بين ٥ر٠ إلى ٨ر٠ وكذلك تتراوح معاملات ثباتها فى نفس المدى . كما أن معاملات الارتباط بين كل منها وبين اختبار صادق للذكاء تختلف فى مدى كبير . وسنختار أكثر هذه الاختبارات شيوعا :

« اختبارات المتروبوليتان للتهيؤ فى القراءة »

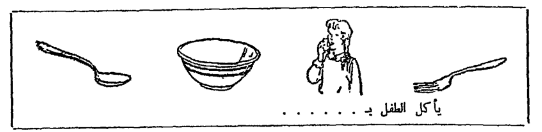
اختبارات المتروبوليتان للتهيؤ فى القراءة : Metropolitan Reading Readiness Test تتضمن ٦ اختبارات هى :

الاختبار الأول : معنى الكلمة word meaning وفى هذا الاختبار ينطق الفاحص بكلمة وعلى المفحوص أن يختار الصورة التى تمثل هذه الكلمة من بين ٤ صور فى كل صف (انظر شكل ١٨) .

(شكل ١٩) عينة من اختبار المتروبوليتان للتهيؤ فى القراءة

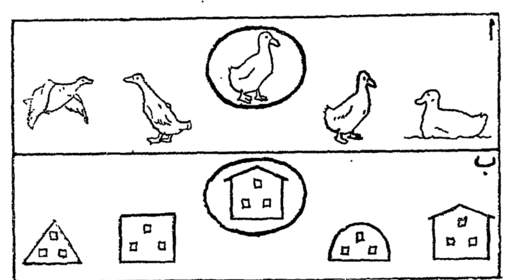


ينام الولد على . . . .

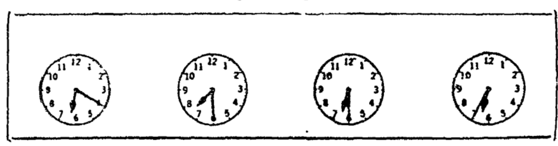


يأكل الطفل بـ . . . .

يقرأ الفاحص الجملة فيشير إلى صورة الشىء الذى يكمل اسمه الجملة



والمفروض هنا أن المفحوص سيضع دائرة حول أحد الأشكال الأربعة الذى يطابق الشكل المحاط بدائرة



علم على الساعة التى تشير إلى السادسة والنصف . . . . . .

الاختبار الثانى : الجمل Sentences : وفيه يختار المفحوص جملة مصورة فى أحد صور من بين الصف الذى يتكون من ٤ صور . وينطق الفاحص الجملة المكتوبة وسط هذا الصف ثم ينتقل إلى جملة تالية من بين ٤ جمل فى الصف الثانى وعلى المفحوص أن يتعرف على الصورة التى تمثلها فيوصل بينهما خطا .

الاختبار الثالث : المعلومات info motion : وفى هذا الاختبار يختار المفحوص صورة من كل صف تؤدى الوظيفة التى يقوم بها الاسم الذى ينطق بها الفاحص . فمثلا يقول الفاحص « اختر الصورة التى تقوم بوظيفة التقاط الصـــورة فيعلم المفحوص على آلة التصوير ويوصل بينها والجملة خطا .

الاختبار الرابع : المزاوجة matching : وفيه يقوم المفحوص بوضع خط يوصل بين صورة فى صف يتكون من ٤ صور وبين صورة أخرى تماثل إحداها . والصورة التى عليه أن يجد مثيلتها محاطة بخط (انظر شكل ١٩ الأوزة) والصـــور الأربع فى الصف تقترب من بعضها فى خصائص وتختلف فى أخرى . فمثلا فى أحد الصفوف بطة وأوزة ودجاجة راقدة وأحد الطيور المنزلية الأخرى ووسط هذه الصور توجد صورة أوزة محاطة بدائرة . ويطلب من المفحوص أن يوصل هـــذه المحاطة بمثيلتها . وهذا يتطلب منه معرفة أوجه الشبه والاختلاف بين الصور ومعرفة الخصائص الهندسية لكل منها أحياناً .

الاختبار الخامس : الأعداد Numbers : وهنا يطلب من المفحوص القيام ببعض العمليات الحسابية البسيطة ومعرفة بعض المفاهيم الكمية . وهذا الاختبار يشبه ما تأتى به اختبارات الاستعدادات الفارقية واختبارات الذكاء العام .

الاختبار السادس : النسخ copying ويقوم المفحوص فيه بنقل أى إعادة رسم بعض الأشكال الهندسية البسيطة والحروف والأرقام . وهذا يتطلب درجة من النمو الجسمى وأخرى من النمو العقلى .

ولهذه البطارية صورتان متكافئتان . ومعاييرها هى الدرجات المئينية وتستخدم كاختبار للتهيؤ فى القدرة وفى الأرقام كما تستخدم كمقياس للاستعداد العام للتعليم .

تكونت عينة تقنينها من ١٥٠٠٠ تلميذاً وكانت تطبق فى أول شهر يدخل فيه التلميذ المدرسة الابتدائية فى السنة الأولى . ويتبع هذا تطبيق اختبار « رسم رجل » Draw a man Test . ويعتبر هـذا الاختبار إضافياً . ويصحح اختبار رسم الرجل فيحصل التلميذ على درجة نضعه حسبها فى فئة من خمسة فئات . ثم يوضع التلميذ حسب درجته فى البطارية فى فئة . ويوجد معامل الارتباط بين الفئة فى البطارية والفئة فى الاختبار ويكون هذا المعامل دليلا على صدق البطارية .

اختبارات جيتس للتهيؤ فى القراءة

صممت هذه الاختبارات لقياس التهيؤ لتعلم القراءة ، وللتنبؤ بسرعة نمو القدرة القرائية ، ولتشخيص مكانة التلميذ وبيان حاجاته فى كل قدرة هامة يحتاجها عند تعلم القراءة .

وتعطى درجة مركبة أو متوسط درجات التلميذ على كل اختبار فرعى لتحديد مرحلة التهيؤ التى وصل إليها هذا التلميذ . وهذه الدرجة المركبة قد تستخدم فى تقرير ما إذا كان التلميذ مستعداً للبدء فى البرنامج المنتظم للدراسة ، ولتقدير نجاحه المحتمل إذا كان قد بدأ القراءة فعلا ، ولتقدير ما هى الأعمال التى يحسن أن يبدأ بها قبل تعلمه القراءة ، وهذه الدرجة

المركبة ، على أى حال تغطينا فقط معلومات عامة عن اســتعداد التلميذ للقراءة .

ويفيد المدرس من هذه الاختبارات فى مقارنة درجات التلميذ الواحد فى مختلف القدرات القرائية وفى مقارنته بغيره فى قدرة معينة . وقد تم إعداد هذه الاختبارات بعد القيام بسلسلة من البحوث التى استمرت أعواماً طويلة . وفى إحدى هذه الدراسات استخدم حوالى ١٠٠ اختبار مختلف ، وامتحانات ، وتقديرات . . الخ . ذلك لتحديد أفضل طرق قياس هذا التهيوء . والقدرات التى كشفت عنها هذه الدراسات وغيرها من الدرسات المماثلة وضعت لها اختبارات من النوع المألوف وجربت على كل تلاميذ إحدى البلدان التى اتبعت طرقاً مختلفة فى التعليم . وبعد دراسة نتائج هذه الاختبارات ، أعدت اختبارات مراجعة وجربت على مجموعات أخرى من التلاميذ بين ســنة ١٩٣٨ وسنة ١٩٣٩ . ويمثل الاختبار الحالى مدى الإفادة من هذه الدراسات . وهو يشمل خمس اختبارات فرعية هى : —

الاختبار الأول : تعليمات الصور Picture Directions فى هــذا الاختبار صور ثلاث الأولى صفحة كاملة عن منظر مزرعة ، والثانية نصف صفحة عن منظر مدينة ، والثالثة نصف صفحة عن المنظر الداخلى لمحل تجارى عام .

يقوم الفاحص بالتحدث إلى التلاميذ عن المواقف أو الأشياء الممثلة فى فى الصورة ويطلب منهم تنفيذ تعليمات معينة بوضع علامات على الصورة وهكذا يمكن قياس القدرات الآتية : —

(أ) القدرة على الإصغاء لما يقوله الفاحص .

(ب) القدرة على فهم ما قاله .

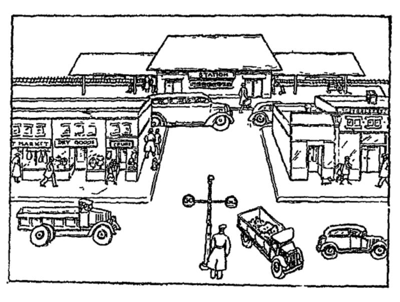
(ج) القدرة على تذكر هذا القول لمدة قصيرة ؛

(د) القدرة على تمثل واستخدام كلمات ومفهومات الحياة اليومية العامة المتنوعة التي تتعلق بالريف ، والمدن ، والمحلات ؛

(هـ) القدرة على تفسير صور يشيع وجودها في الكتب الابتدائية ؛ وهــذا يتضمن فهم الصورة الكلية ، ووضع الأشياء فيها ، ؛ ؛ الخ .

(و) القدرة على استخدام كل القدرات السابقة في تنفيذ التعليمات ؛

ومادام هذا الاختبار يتطلب أن يصغي التلميذ إلى ما يقوله المدرس وأن ينفذ التعليمات لمدة زمنية متوسطة الطول ، فالاختبار إذن يقيس الانتباه في نشاط يشابه ما يحدث في أثناء الدرس الشفوي المعتاد في المدرسة ؛



(الشكل ٢٠)

فقرة من اختبار جيتس للتهيؤ في القراءة وهي الفقرة الثانية من اختبار تعليمات الصور

وهو يقيس ، بالتالى ، عدداً من القدرات القاعدية على تعلم القراءة فى موقف الفصل . والقدرة التى يقيسها هذا الاختبار هى محصلة أوجه النشاط التى يخبرها التلميذ قبل أن يدخل المدرسة إطلاقا . وهذه الخبرات تتعلق بكل أنواع الصور ، والإصغاء إلى كل أنواع التعليمات ، والجمل ، والقصص ، والتعليمات ، وتنفيذ كل أنواع التعليمات فى البيت أو مدرسة رياض الأطفال .

الاختبار الثانى : مزاوجة الكلمات Word Matching : هذا الاختبار مصمم لقياس مدى تعرف التلميذ على الكلمة المطبوعة سواء كان يستطيع قراءتها أم لا يستطيع . والكلمات كلها شائعة فى كتب الأطفال ، وعناوين الجرائد ، والمجلات ، والمحلات .

وفى الاختبار ١٨ مربعا فى كل ٣ كلمات منها كلمة مكررة أى أن المربع يشمل ٤ كلمات . ويضع المفحوص علامة بين الكلمة المكررة مرتين .

الاختبار الثالث : المزاوجه بين الكلمة والبطاقة word-Card Matchiny وفى هذا الاختبار تقدم للتلميذ عدة بطاقات، على كل منها كلمة يكتبها المدرس بخط معتدل كبير جداً والحروف ذات حجم مقنن وعدد البطاقات ٢٠ بطاقة . ويطلب الفاحص من التليمذ أن يضع علامة على الكلمة التى تشابه الكلمة المطبوعة فى البطاقة من بين ٤ كلمات فى كل مستطيل مطبوع فى كراسة التعليمات وعدد المستطيلات ٢٠ مستطيلا .

الاختبار الرابع : السجع : Rhyiming وفى هذا الاختبار صفحتان كل منهما تشمل ٧ مجموعات من الكلمات . وكل مجموعة تتكون من ٤ كلمات . وفيه يساعد الفاحص التلاميذ على نطق كلمات كل مجموعة ثم يطلب منهم وضع علامة على كلمة ، فى كل مجموعة ، تتفق مع كلمة أخرى مكتوبة أمام المجموعة ينطقها الفاحص مثل نام ، لعب ، يمشى ، أصبح . وأمام

هذه الكلمات كلمة مثل قام وعلى المفحوص أن يضع علامة تحت أى من الأربعة التى تتفق معها (والإجابة الصحيحة هى « نام ») .

الاختبار الخامس : قراءة الحروف والأرقام Reading Letter and Number : وهذا الاختبار يتكون من ثلاثة أجزاء . الجزء الأول حروف مكتوبة بخط كبير capital والثانى مكتوبة بحروف صغيرة small والجزء الثالث الأرقام من صفر إلى ٩ . والحروف ، كبيرة كانت أو صغيرة ، والأرقام غير مرتبة . وينطق المدرس بالحرف أو بالرقم وعلى المفحوص أن يشير إليه فى كراسة التعليمات :

ومن الطبيعى عند نقل هذه الاختبارات إلى العربية أو إعدادها كى تناسب ظروفنا المختلفة يجب ، فضلا : عن عمليات التقنين والصدق والثبات أن تراعى قواعد اللغة العربية كأن تستبدل حروف النسخ والرقعة بالحروف الكبيرة والصغيرة وأن تحل صور مألوفة فى بلادنا محل صور غريبة فى ثقافتنا مألوفة فى بلادها الأصلية :

والمعيار المستخدم هو الدرجات المئينية ، ومعاملات الثبات مرتفعه جداً ؛ ومعاملات الصدق بالارتباط مع اختبارات صادقة عالية أيضاً : إذ أن متوسطها ٧٦ر٠ :

اختبارات مسح القراءة

تقيس هذه الاختبارات فى كل مستويات التعليم المستوى العام الذى وصل إليه المفحوص فى نموه القرائى وهى تستخدم لتصفية التلاميذ الذين يمكن مساعدتهم بواسطة القراءة العلاجية ، وللتنبؤ بالنجاح فى المناهج الدراسية . وللتأكد مما إذا كانت القدرة القرائية لدى المفحوص ضعيفة وأن هذا هو السبب الذى من أجله كان أداؤه ضعيف فى المدرسة أو فى اختبار عقلى جمعى :

تشمل القدرة القرائية جانبين . هما السرعة والفهم وبالتالى ، ينجح

اختبار مسح القراءة إذا وضع فى الاعتبار تدخل كل من هذين العاملين معاً . وقد حاول بعض علماء القياس اختبار كل عامل على حدة ولكنهم لم ينجحوا فى هذا . وليس ذلك بغريب علينا إذ أنه يرجع مبدأ عام متفق عليه فى علم النفس عموما والقياس خصوصا وهو أن الكل المركب لا يحسن تفتيته إلى مكوناته وقياس كل على حدة :

ويمكن ، من حيث النظـرية ، عزل عاملى السرعة والفهم Comprehension وقياس كل منهما على حدة بتثبيت أحدهما وتقدير الآخر . فيمكن مثلا إلغاء أثر السرعة أو تقليل أهميتها ذلك بإعطاء وقت كاف للإجابة عن الاختبار . وبالتالى تمثل درجة المفحوص قدرته على الفهم . أما عند تثبيت عامل الفهم بوضع اختبار يتوقع أن جميع المفحوصين يفهمون مواده ، لقياس عامل السرعة ، فهو أمر أصعب من سابقه . إذ أن لكل فرد عدد كبير من السرعات القرائية ، وتختلف سرعة القراءة باختلاف الهدف من القراءة ونوع ومستوى وصعوبة وطبيعة المادة المقروءة . ويميل مجيدو القراءة إلى تخفيض سرعتها كلما زادت صعوبة المادة التى يقرأونها .

المشكلة بالنسبة للفاحص إذن هى : كيف يمكن وضع قطعة يقرأها كل التلاميذ بنفس الدرجة من التركيز وبنفس الجهد للفهم ؟ هذه المشكلة يواجهها المدرس عندما يريد قياس سرعة القراءة . مما يزيد الأمر تعقيداً أنه يمكن أن تزيد السرعة ويقل التركيز والفهم فيقل التحصيل . وإذا لم يضبط الجهد المبذول فى الفهم ، بأن يطلب من المفحوص أن يجيب عن أسئلة تتعلق بما قرأه ، أو أن يكتشف السخافات أثناء القراءة ، كانت النتيجة هى غموض المادة التى يقرأها التلميذ بسرعة حتى يحصل على درجة عالية فى أحد اختبارات سرعة القراءة .

ومن هنا يتبين لنا أنه يصعب فصل سرعة القراءة عن مستوى الفهم وهذه المشكلة تتعرض لها اختبارات المسح والتشخيص . ووجه الصعوبة ، كما قلنا ، هو أن سرعة القراءة تعتمد على صعوبة المادة المقروءة وسبب قراءة المفحوص لها . ومن هنا أيضاً تكون للمفحوص عدة سرعات فى القراءة تختلف من اختبار إلى آخر وقد تختلف بين فقرات الاختبار الواحد . كما أن سرعة القراءة تتوقف على خبرات المفحوص الخاصة .

وعلى هذا فاختبارات مسح القراءة ، وخاصة عند استخدامها فى المدرسة الابتدائية والإعدادية ، تتبع اختبارات التصنيف العام . ولكنها ، عند استخدامها فى مستوى المدرسة الثانوية والجامعة ، تصبح اختبارات خاصة لأنها تقيس مجالات خاصة مثل العلوم والجغرافيا والتاريخ والأدب .

وبهذه الطريقـــة يمكننا أن نتبين أن اختبارات المسح تواجه الصعوبات التالية :

١ ــ تتداخل الدرجات على السرعة مع الدرجات على الفهم بحيث يستطيع التلميذ أن يرفع إحداهما على حساب الأخرى ، وهو قد يفعل هذا دون قصد .

٢ ــ تكون الاختبارات التى يفترض أنها تقيس الفهم ، غالباً ، مشحونة بعامل السرعة : ومثل هذه الاختبارات محدودة القيمة إذا أردنا بها التشخيص ، ولكنها تعتبر أدوات تنبؤية جيدة .

٣ ــ يغطى اختبار القراءة مجالا محدوداً مختاراً من محتوى المواد الدراسية ، وبما أن القدرة القرائية تختلف بعض الشيء باختلاف المادة ، فإننا نجد نتيجة لهذا السبب ونتيجة للطريقة التى اتبعت فى اختيار فقرات الاختبار ، أن القيمة التنبؤية باختبار مسح القراءة تتحدد بمجال المادة الدراسية التى اختيرت فقراته منها . وعلى هذا لا يصح استخدام اختبار

لمسح القراءة فى مادة كالتاريخ كأداة للتنبؤ أو التشخيص فى مادة أخرى كالحساب ولا حتى فى مادة قريبة من التاريخ مثل الجغرافيا .

٤ ــ تختلف اختبارات مسح القراءة فى مستويات الفهم التى تقيسها . فبعض الاختبارات يتطلب فهما بسيطاً عاماً ، بينما يتطلب فى البعض الآخر تفسيراً عميقاً تاماً شاملا . ولكننا إذا أردنا أن نتعرف على القارئ الماهر لا نكتفى بفهمه الجمل ، أو هذا ما يجب ، بل وأيضاً باستخلاصه الفكرة العامة فى القطعة التى قرأها ، وبجمعه الأفكار من الجمل المتفرقة ، وبمتابعته مناقشة منطقية تستغرق أحيانا صفحات . ولكن اختبار مسح القراءة الواحد لا يمكنه أن يقيس كل أنواع ومستويات الفهم منفصلة بنفس الكفاءة .

تشخيص القراءة

الاختبار التحصيلى التشخيصى ، فى أحسن صورة ، هو أداة تأثيرية . فعلى المدرسين ، مستعينين بهذه الاختبارات أو بغيرها أو مستخدمين طرقهم الأخرى الخاصة ، أن يدرسوا كل تلاميذهم ليحددوا أسباب مشكلاتهم . والاختبار التشخيصى يحاول أن يزيد كفاءة هذه الدراسة فيضع أمام ناظرى المدرس كل فهم يمكن الوصول إليه من بحث أسباب ومصادر الفشل فى عمليات التعلم . والاختبار التحصيلى التشخيصى الجيد يوجه انتباه المدرس إلى كل أو أغلب وأهم نواحى العملية القرائية التى يجد فيها التلميذ صعوبة . ومثل هذه الجهود يجب أن تستند إلى بحوث شاملة لتحديد أنماط الأخطاء التى يغلب أن يقع فيها التلاميذ ، والتى يغلب أن يعانى منها معظم أفراد كل فئة مرضية . ثم توضع هذه الأخطاء فى قوائم تشخيصية .

وهناك اختبارات تشخيصية أبسط من الأنواع السابقة . وبعضها يجرى جمعياً . وهى تحوى اختبارات فرعية لقياس مختلف أنماط القدرة القرائية .

ويمثل الأداء بصفحات نفســية تبين المستويات النسبية لقدرات الفرد القرائية . ومنها يمكننا تكشف نواحى الضعف والقوة لديه . وأشهر هذه الاختبارات :

اختبار أيوا للقراءة الصامتة

يشمل اختبار أيوا للقراءة الصامتة The Iwa Silent Reading Test الاختبارات الفرعية التسع الآتية :

الاختبار الأول والثانى : سرعة القراءة : يقرأ التلميذ مقالتين قصيرتين ، ثم يجيب عن أسئلة بسيطة عن القطعتين . وتحسب للفرد درجتان إحداهما للسرعة والأخرى للدقة . ولأنه تحسب درجتان ، عد الاختبار اختبارين .

الاختبار الثالث : القراءة الموجهة : directed Reading يجيب التلميذ عن أسئلة عن الفقرات بالإشارة إلى مكان الإجابة عن السؤال بين كلمات الفقرة بوضع خط تحتها Underlining . وهذا الاختبار ، كالإختبارات الفرعية التالية ، له توقيت منفصل لا يكفى لإنهاء الاختبار غالباً .

الاختبار الرابع : الشعر : يقرأ التلميذ قطعا من الشعر ويجيب عن أسئلة عن معناها .

الاختبار الخامس : معنى الكلمة : تعرض ٤ قوائم من الكلمات فى العلوم الاجتماعية ، والعلوم ، والرياضيات ، واللغة . وبجوار كل كلمة خمسة متغيرات يختار الصحيحة منها كمعنى للكلمة . ويتم إجراء هذا الاختبار فى ٤ مراحل كل مرحلة من ٩٠ ثانية . وعند انتهاء هذه المدة ينتقل التلميذ إلى المرحلة التالية . وتوضع درجة عن أداء كل جزء .

الاختبار السادس : معنى الجملة : يقرأ التلميذ جملا تعالج معنى كلمات وتتضمن معلومات بسيطة . وكل جملة هى حقيقة أو معنى تفسيرى لكلمة ويجيب المفحوص بصحيح أو خطأ عن كل .

الاختبار السابع : معنى الفقرة : Paragraph meaning يقرأ التلميذ فقرات ويجيب أسئلة عن كل . وهذه الإجابة تطلب منه إعمال الذاكرة المباشرة والتعرف على الفكرة الرئيسية .

الاختبار الثامن : استخدام الفهرست : Use of index يعرض على التلميذ فهرست موجز ويطلب منه أن يحدد صفحات الكتاب التى يعتقد أنها تتضمن إجابات عن أسئلة معينة .

الاختبار التاسع : الكلمات الرئيسية Key Words : يشير التلميذ إلى الكلمات التى يجب أن يبحث تحتها فى الفهرست كى يجد الإجابة عن سؤال مثل : « ما هى قيمة إنتاجنا السنوى من المواد المعدنية » ؟

وتختلف اختبارات التشخيص فى القراءة من حيث درجة تحليل كل اختبار للقدرة القرائية إلى عدد من القدرات الفرعية . فمن جهة نجد الاختبارات الجمعية وفيها يحصل التلميذ على ٣ أو ٤ درجات منفصلة كل عن إحدى الجوانب المقاسة ، ومن أخرى نجد الاختبارات الفردية التى توضح الأخطاء النوعية التى يقع فيها بعض المفحوصين . وتشمل الاختبارات الأخيرة قوائم مقننة لتلك الأخطاء . ويستخدم بعضها جهاز التاكستسكوب لتتبع سرعة حركات العين أثناء القراءة . وهذا الجهاز ، ببساطة ، يمكن المجرب من وضع البطاقة فيه ويجذب أو يدار محرك فيستمر عرض البطاقة مدة يتحكم فيها الفاحص . وأثناء هذه المدة يستطيع المفحوص أن يقرأ المكتوب على البطاقة إن كان يقرأ بالطريقة الكلية . وهناك أزمنة

مقننة لمتوسط المدة التى تستغرقها قراءة الكلمة . فإذا زاد عنها التلميذ كان يقرأ بالطريقة الجزئية .

ومن أهم أمثلة اختبارات القراءة التشخيصية الجمعية فى المدارس الابتدائية والإعدادية اختبارات جيتس للقراءة الأساسية Gates Basic Reading Tests وهى تقيس السنوات الدراسية من الثالثة الابتدائية حتى قرب نهاية المدرسة الإعدادية ؛ وهو جزء من سلسلة اختبارات لها شهرتها العالمية . ويشمل ٤ مقاييس منفصلة هى : اختبار فهم المعنى العام — اختبار التنبؤ بنتائج بعض الأحداث المعينة — اختبار تحديد اتجاهات معينة — اختبار ملاحظة التفاصيل . ولهذه السلسلة من الاختبارات ٤ صور متكافئة .

ويمكن اعتبار اختبارى « أيوا » « وجيتس » اختبارات للمسح ، بل إنها أيضاً تشمل اختبارات مسح للقراءة . ومعاملات الارتباط بين الأجزاء الأربعة فى اختبار « جيتس » مرتفعة . ومعاملات ثبات كل منها مرتفعة أيضاً . أما معاملات ثبات اختبارات « أيوا » فهى منخفضة ، وإن كانت معاملات الارتباط الداخلية بينها منخفضة أيضاً ، مما يجعلها أداة تشخيصية لا يمكن الاعتماد عليها تماما . كما أن درجات الأفراد عن اختبارى السرعة فى « أيوا » قليلة الدلالة ، للأسباب التى ذكرناها .

ولكن هذه الانتقادات توجه إلى كل اختبارات القراءة ، فضلا عن نقد أساسى آخر هو أنها لا تقيس المهارات الأولية المطلوبة فى الأعمال المكتبية . وإن كانت الدرجات عن الاختبارات الفرعية منخفضة الثبات ، إلا أن معاملات ثبات الدرجات الكلية عن الاختبارات المتضمنة فى كل بطارية منها معاملات مرتفعة وهذه ميزة فضلا عن ميزة أخرى هامة هى أن معايير هذه الاختبارات حسبت من عينات كبيرة وممثلة من التلاميذ . .

وأما بطارية امتحانات مونرو للقراءة التشخيصية Monroe Diagnostic Reading Examination فهى تعتبر برنامجا للاختبار أكثر منها بطارية

مفردة : ولإكمال هذا البرنامج يحتاج الأمر عدة ساعات وتتطلب تطبيق اختبار للذكاء ، وعدة اختبارات مقننة للمسح فى القراءة ، وبعض الاختبارات التحصيلية الأخرى ، وبعض اختبارات القراءة الخاصة . وللبطارية جزء إضافى للتحليل الكيفى لصعوبات القراءة وللبيانات المفصلة عن الأخطاء :

والبطارية الرابعة هنا هى بطارية دوريل لتحليل صعوبات القراءة Durell Analysis of Reading Difficulty Battery فهى أقصر من السابقة وتطبق من السنوات الأولى حتى السادسة الابتدائية . وقد تم تقنين ما أسفرت عنه بحوث التحليل المفصل لصعوبات القراءة الشفوية والتذكر والاسترجاع والتعرف على الكلمات والنطق والكتابة والهجاية . وهى تستخدم جهاز التاكستسكوب لمقارنة سرعة التعرف على الكلمات بقراءة الكلمة فى ظروف معينة ثم قراءتها فى أخرى بسرعة أخرى وهكذا . وتعطى درجة جزئية عن تصحيح المفحوص لقائمة من الكلمات الخاطئة بعد أن يكون قد فهم قائمة بالكلمات الصحيحة . وقننت هذه البطارية على ٤٠٠٠ تلميذ .

ويجب أن نخص بالذكر اختبارات القراءة التشخيصية Diagnostic Reading Test التى قامت بها جمعية الاختبارات التشخيصية الأمريكية . وهى من جزئين أحدهما مسحى والآخر تشخيصى . واختبارات الجزئين كلها جمعية فيما عدا اختبار واحد فى الجزء التشخيصى هو اختبار القراءة الشفوية . وتستخدم هذه البطارية ابتداء من أول المدرسة الإعدادية حتى السنة الأولى بالجامعة .

وللجزء المسحى ثمانية صور مكافئة . ويحصل الفرد فيه على ٣ درجات منفصلة لأن البطارية تقيس ٣ مجالات هى : سرعة القراءة مع الفهم المعقول للمادة المقروءة — المفردات العامة — القدرة على فهم الكتب .

المدرسية . وقد أوجدت معايير للفرقة الدراسية بناء على الدرجة المركبة عن البطارية ككل . كما يعد التلميذ فى فرقة دراسية معينة بناء على الدرجات عن الاختبارات الفرعية الثلاث . والعينة التقنينية أخذت من عدة مناطق جغرافية ومدن ومدراس عامة وخاصة وكثير من المعاهد والكليات .

والقسم التشخيصي يتطلب وقتاً أطول من المسحى ، ويمكن أن يطبق بصورة جمعية فى الفصول . ويمكن الاكتفاء بتطبيق الأجزاء من البطارية التشخيصية التى ظهرت صعوبات القراءة فيها أثناء تطبيق الجزء المسحى . ويطبق الجزء المسحى التشخيصى فى المجالات الثلاث السابقة بالإضافة إلى اختبار لمهارة التعرف على الكلمات . وتتضمن البطارية التشخيصية مفردات فى قواعـــد اللغة ، والأدب ، والعلوم ، والدراسات الاجتماعية ، والرياضيات ، وفهم مواد من الكتب الدراسية يقرأها المفحوص أو تقرأ له . وتحسب سرعة القراءة للمواد المختلفه وعن ربط الكلمات سواء كان ذلك شفوياً أم فى صمت .

وآخر هذه الاختبارات يتضمن قوائم للكلمات ، وأخرى للفقرات paragraph . والجزء الشفوى يتأثر بعوامل مثل مهارة استخدام المفردات ، والأخطاء فى التعرف على الكلمات ، وطريقة ربط كلمات غير معروفة . وهناك جزء للملاحظات الكيفية ، وقائمة بالأخطاء الشائعة ، وربط الكلمات فى صمت . وهذا الجزء يعطى درجة للقدرة على تقييم الكلمة ، وأخرى للقدرة على سماع وتمييز الكلمات .

والمعيار هو الدرجات المئينية . ومعاملات ثبات القسمين أكثر من ٨٠ر٠ ويجب ، عند الحديث عن الاختبارات التشخيصية فى القراءة ، أن نتحدث عن :

أولا : بناء هذه الاختبارات .

ثانياً : تفسيرها واستخدامها .

والجزء الثانى من الحديث موجه إلى الأخصائيين النفسيين الذين يستخدمونها فردياً أو فى الفصول . فإن تشخيص صعوبات القراءة ، وطريقة علاجها مسألتان تحتاجان إلى أخصائيين مدربين . إذ أن تشخيص عجز القدرة القرائية يعتمد على الطريقة الإكلينيكية فى دراسة الحالات ، وإلى جمع معلومات إضافية عن الإمكانيات الحسية ، والنمو الحركى ، ومعلومات عن الصحة ، والعلاج ، والتعليم ، والمنزل والأسرة ، والنواحى الانفعالية عن الفرد . ذلك أن تأخر القدرة القرائية قد يكون عرضا لاضطراب الشخصية . وبما أن الاختبارات الجمعية للمسح والتشخيص تتطلب الاهتمام بالأفراد ، مما يجعلنا نتهى إلى القول بأن تقييم نقاط الضعف والقوة فى القدرة القرائية ، يدخل فى صميم الخدمات الإكلينيكية .

الرياضيات

اهتم الكثيرون بتشخيص العجز والتأخر فى القدرات الرياضية اهتمامهم باختبارات القراءة ، إبتداء من مستوى المدرسة الإبتدائية حتى أعلى مستويات التعليم . ومع أن اختبارات الرياضيات ليست بنفس الكثرة والشيوع ، إلا أنها تنقسم أيضاً إلى الأنماط الثلاث السابقة وهى :

أولا : اختبارات التهيؤ فى الرياضيات .

ثانياً : اختبارات مسح الرياضيات .

ثالثاً : اختبارات التشخيص فى الرياضيات .

ويغلب أن تجمع الاختبارات بين الفئتين الأولى والثانية فى بطارية واحدة .

التهيؤ والتنبؤ

نجد أن بطارية المتروبوليتان ، التى سبق عرضها ، تعطى درجة للاستعداد للقراءة عن اختبار واحد للأرقام . ومن ثم فهى ليست ذات دلالة كبيرة . وإن كانت تشير إلى الاستعداد للنجاح المدرسى خاصة إذا طبقت على أطفال السنة الأولى الابتدائية . وتقوم اختبارات التهيؤ بنفس الوظيفة فى المستويات الأعلى حين تقدم اختبارات فى الجبر والهندسة .

ومن أهم هذه الاختبارات ، اختبار « لى » لقدرة الجبر Lee Test of Algebra Ability وهى تتكون من ٤ اختبارات فرعية للمسائل الحسابية ـ ومقارنة الأرقام ـ وسلاسل الأعداد ـ والمعادلات . أما اختبار أورليانز للتنبؤ بالجبر Orleans Algebra Prognosis Test فهو يتكون من ١٢ جزءاً منها ١١ جزءاً عن دروس بسيطة فى الجبر متبوعة بأسئلة عن محتواها . والجزء الأخير هو تلخيص لما تعلمه التلاميذ من المواد الجديدة .

اختبار « لى » إذن يقيس التهيؤ أكثر مما يركز على التنبؤ ؛ بينما اختبار « أورليانز » يقوم المفحوص فيه بعمليات بسيطة تهتم بالتنبؤ أكثر . وقد حسبت معاملات صدق كل من الاختبارين بإيجاد معاملات الارتباط بين الاختبار ودرجات الفترات وبين درجات الاختبار ودرجات نفس التلاميذ فى اختبارات تحصيلية أخرى صادقة .

المسح

اختبارات مسح الرياضة فى المدرسة الابتدائية تقتصر على الحساب وفى المستويات الدراسية الأعلى تشمل أغلب فروع الرياضة . وهى تكون جزءا هاما من البطاريات التعليمية . وتطبق هذه الاختبارات فى مناهج الرياضة المختلفة مثل الحساب والجبر والهندسة . ومن أمثلتها اختبار أيوا للمهارات

الأساسية عند كل تلميذ Iwa Every Pupil Test of Basic skills ونجد في المراجعة الحديثة له عدة اختبارات عن المهارات الحسابية تقيس :

١ - المفردات الفنية والمعلومات الكمية الأساسية .

٢ - المهارات الحسابية .

٣ - حل مسائل تقرأ للمفحوص .

ويعطى للتلاميذ في السنوات من الثالثة الابتدائية حتى الشهادة الإعدادية

التشخيص

من أمثلة هذه الاختبارات ، اختبار تشخيص العمليات الرئيسية في الحساب The Diagnostic Test for fundamental processes in Arithmatic وهو فردى . والمسائل تحل شفويا . ولهذا يسهل فيه أن يتناول الطفل المشاكل والعمليات المختلفة . وهناك قائمة لحساب الأخطاء تتضمن الأخطاء الحسابية الشائعة . اختيرت مسائلها بحيث تتضمن صعوبات توقع بالمفحوص في هذه الأخطاء . ويطبق هذا الاختبار من السنة الثانية الابتدائية حتى أواخر المرحلة الإعدادية تقريبا . ولم يعد له توقيت ولا معايير ولا تعطى درجة كلية . والمسائل ، وقائمة مراجعة الأخطاء يعتمدان في تصحيحهما على التحليل الكيفي أكثر من اعتمادهما على التحليل الكمي .

اختبارات التشخيص في الحساب

ومن أشهر الاختبارات هنا ، اختبارات التشخيص في الحساب Diagnostic Tests in Arithmetic التي أعدها مونرو Monroe ويتكون من أربعة سلاسل من الاختبارات هي :

السلسلة الأولى : وتشمل عمليات أرقام صحيحة integers وتحوى الاختبارات الآتية :

١ ــ اختبار جمع ٣ أعداد . وكل عدد يتكون من رقم واحد . ويشمل ٢٤ عملية . وزمنه نصف دقيقة .

٢ ــ اختبار طرح عدد من رقم واحد من أعداد من رقمين ويشمل ٢٤ عملية . وزمنه نصف دقيقة .

٣ ــ اختبار ضرب عدد من رقم واحد فى عدد من أربعة أرقام ويشمل ٢٠ عملية . وزمنه دقيقة واحدة .

٤ ــ اختبار قسمة عدد من ٤ أرقام على عدد من رقم واحد ويشمل ٢٤ عملية . وزمنة دقيقة واحدة .

٥ ــ اختبار جمع ٥ أعداد ، كل عدد من ٤ أرقام ، رأسيا . ويشمل ١٤ عملية . وزمنه ٣ دقيقة .

٦ ــ اختبار قسمة عدد من ٤ أرقام على عدد من رقمين ويشمل ٢٤ عملية . وزمنه ٢ دقيقة .

السلسلة الثانية : وتشمل عمليات أرقام صحيحة . وتحوى الاختبارات الآتيــة .

٧ ــ اختبار جمع ١٣ عدداً . وكل عدد من رقم واحد ويشمل ١٥ عملية وزمنه ٢ دقيقة .

٨ ــ اختبار ضرب عدد من رقمين فى عدد من ٤ أرقام ويشمل ١٢ عملية وزمنه ٣ دقيقة .

٩ ــ اختبار طرح عدد من ٣ أرقام من عدد من ٣ أو ٤ أرقام ويشمل ٢٤ عملية وزمنه دقيقة واحدة .

١٠ ــ اختبار ضرب عدد من رقمين أو ثلاثة أرقام فى عدد من ثلاثة أرقام ويشمل ٢٤ عملية وزمنه ٢ دقيقة .

١١ – اختبار قسمة عدد من ٥ أرقام على عدد من رقمين . ويشمل ١٢ عملية . وزمنه ٤ دقائق .

السلسلة الثالثة : وتشمل عمليات على الكسور الاعتيادية وهى تحوى الاختبارات الآتية :

١٢ – اختبار جمع كسرين اعتياديين البسط رقم واحد والمقام رقم أو اثنين . ويحوى ١٥ عملية وزمنه دقيقة ونصف .

١٣ – اختبار طرح ، بنفس الطريقة السابقة يحوى ١٥ عملية وزمنه ٢ دقيقة .

١٤ – اختبار ضرب ، بنفس الطريقة السابقة يحوى ١٥ عملية وزمنه دقيقة واحدة .

١٥ – اختبار جمع كسرين اعتياديين ، بنفس الطريقة السابقة . ولكن فى الاختبارين الماضيين يسهل على التلميذ إيجاد مقام هو المضاعف البسيط للمقامات المختلفة ، بينما فى هذا الاختبار والاختبار المشترك التالى لايسهل هذا . ويتكون هذا الاختبار من ١٥ عملية وزمنه ٢ دقيقة .

١٦ – اختبار قسمة كسر اعتيادى بسيط على آخر ويشمل ١٥ عملية وزمنه ٢ دقيقة .

السلسلة الرابعة : وتشمل عمليات ضرب وقسمة كسور عشرية . وبها الاختبارات الآتية :

١٧ – اختبار قسمة رقم صحيح وكسر عشرى على كسر عشرى وأمام عملية القسمة الإجابة الصحيحة دون وضع العلامة العشرية والمطلوب وضع هذه العلامة حسب ما يراه التلميذ صحيحاً . ويشمل الاختبار ٢٤ عملية وزمنه نصف دقيقة .

١٨ – اختبار ضرب رقم وكسر عشرى فى كسر عشرى . ونتيجة عملية

الضرب مرصودة وعلى التلميذ أن يضع العلامة العشرية فى حاصل الضرب . ويتكون من ٣٠ عملية وزمنه نصف دقيقة .

١٩ — اختبار قسمة ، بنفس طريقة اختبار ١٧ ، ولكن المقسوم كسراً من ثلاثة أرقام عشرية . ويشمل ٢٤ عملية . وزمنه نصف دقيقة :

٢٠ — اختبار ضرب رقم وكسر عشرى فى رقم وكسر عشرى ، بنفس طريقة اختبار ١٨ ولكن اختبار ٢٠ أصعب ، ويحوى ٣٠ عملية . وزمنه نصف دقيقة .

٢١ — اختبار قسمة عدد صحيح أو عدد وكسر عشرى على كسر عشرى والنتيجة مكتوبة وعلى المفحوص أن يضع العلامة العشرية فى خارج القسمة . ويشمل ٢٠ عملية . وزمنه نصف دقيقة .

وهى اختبارات سرعة غالباً ، كما يتضح من الزمن المحدد لإجرائها . ويحسب عدد المسائل التى حاول المفحوص الإجابة عنها وعدد الصحيح منها . وهكذا تحسب للتلميذ درجة .

اختبار ستانفورد الجديد فى الحساب

قام الأستاذ إسماعيل القبانى والمؤلف بإعداد اختبارات ستانفورد الجديدة فى الحساب ، وقد حدث ذلك سنة ١٩٣٦ . ومنذ ذلك الوقت طبع الاختبار خمس طبعات وأعدت له معايير محلية . ولكن لم يتم إعداد معايير جديدة حسب التغييرات التى طرأت أخيراً على تنظيم التعليم فى بلادنا .

والاختبار يتكون من جزئين واحد للعمليات الحسابية ، والآخر للتفكير الحسابى . وله صورتان متكافئتان هما الصورة م ، والصورة ن . وإليك تعليمات الاختبار :

كلف التلاميذ باستبعاد كل ما قد يوجد أمامهم . وتأكد أن مع كل منهم قلماً رصاصاً مبرياً برياً جيداً وبعد ذلك قل للتلاميذ « أنا النهارده

حاعطيكم اختبار فى الحساب : حاوزع عليكم كراسات للاختبار ولكن ما حدش منكم يفتحها إلا لما أقول لكم ، علشان تبتــدوا كلكم مع بعض . . . » .

ثم كلف تلميذاً أو تلميذين بتوزيع الكراسات تحت إشرافك وبعد ذلك . قل « اكتبوا البيانات اللى على الصفحة الأولى . . . قدام اسم التلميذ اكتب اسمك . (وقفة) وقدام عمره اكتب عمرك بالسنة والشهر إذا كنت تعرفه . (وقفة) وقدام المدرسة اكتب اسم مدرستك . . . مدرسة كذا . . (وقفة) وقدام السنة الدراسية اكتب السنة والفصل اللى انت فيه . . سنة كذا . . فصل كذا . . (وقفة) وقدام التاريخ اكتب تاريخ النهارده . . كذا . . » .

وبعد أن يفرغوا من ذلك اقرأ عليهم التعليمات الموضوعة فى أسفل الصفحة الأولى من الكراسة بتأن وبصوت واضح بينما هم يتتبعونها معك فى كراساتهم ثم قل « دلوقت انتبهوا كويس واتبعوا التعليمات بالضبط . . ساعة ما أقول ابتدى تفتحوا الكراريس وتبتدوا تجاوبوا . . وساعة ما أقول ضع القلم لازم تبطلوا الكتابة طوالى . . وبعد ما تبتدوا مش عاوز حد يسأل أى سؤال أو يتكلم مع اللى جنبه . . وإذا حد انكسر قلمه يرفع صباعه من غير كلام وأنا أديله قلم غيره . مفهوم . . . لاحظوا أن الاختبار فيه عدة صفحات وكلما تخلصوا صفحة اقلبوها وكملوا على الصفحة اللى بعدها لغاية ما أقول لكم ضعوا القلم (وقفة) . . . يالله دلوقتى . . . ابتدى . . . وهنا لاحظ الوقت فى ساعتك بالضبط وسجله . . . وفى نهاية الزمن المقرر قل « ضع القلم » وتحقق من أنهم جميعاً أطاعوا فوراً . . . ثم اجمع الكراسات .

أما الزمن المقرر فهو :

للفرقتين الأولى والثانية الابتدائيتين : ١٠ دقائق لاختبار العمليات ومثلها لاختبار المسائل .

للفرق التي تلي ذلك[1] : ٣٠ دقيقة لاختبار العمليات ، ٢٠ دقيقة لاختبار المسائل .

تعليمات لتصحيح الاختبار

١ — ضع أمام كل سؤال علامة ✓ إذ كان صحيحاً أو علامة × إذا كان خاطئاً ولا تضع علامة أمام السؤال الذي تركه التلميذ بدون إجابة .

ودرجة التلميذ هي عدد المسائل التي أجاب عليها إجابة صحيحة .

٢ — لا تعطى كسوراً في الدرجات وإذا لم يكتب التلميذ الجواب الصحيح بالضبط فتعتبر الإجابة خطأ .

٣ — إذا أعطى التلميذ جوابين لسؤال واحد تعتبر الإجابة خطأ . ولكن إذا شطب الجواب الأول أو صححه بشكل واضح فيقبل الجواب الثاني .

٤ — إذا وضع التلميذ الجواب في غير محله فأقبله إذا كان واضحاً أنه يقصد أن هذا هو جواب المسألة المطلوبة .

٥ — لا يؤاخذ التلميذ على عدم كتابة التمييز في الجواب .

٦ — إذا كتب التلميذ الجواب على صورة كسر عشري بدل كسر اعتيادي يعتبر صحيحاً .

٧ — إذا كتب التلميذ الجواب على صورة كسر اعتيادي فيجب أن يكون مختصراً إلى أبسط صورة وإلا يحسب خطأ .

أولا : اختبار التفكير الحسابي الصورة م .

يتكون من ٤٠ فقرة مثل :

١ — محمد عنده ٦ أرانب بيضاء و ٥ أرانب سوداء . فكم أرنباً عنده ؟

(الجواب ....)

---

(١) توازي الفرقة الأولى والثانية (نظام قديم) الفرقة الرابعة والخامسة (نظام حديث) .

٤٠ – اشترى رجل بلحاً بسعر كل ٦ أرطال ثمانية قروش . وباعه بسعر كل ٣ أرطال ثمانية قروش . وكان ربحه من ذلك ٦ر١ من الجنيهات . فكم رطلا من البلح باعها ؟ (الجواب ————————)

وتناظر هاتين الفقرتين الفقرتان التاليتان فى الصورة : ن

١ – إذا وضعت ٥ طيور و ٤ طيوراً معاً ، فكم طيراً تصير ؟

(الجواب ————————)

٤٠ – منزل ثمنه ٥٠٠٠ جنيه ، ويدفع صاحبه عليه ضريبة سنوية قدرها ٦٠ جنيها ، ويصرف عليه ٢٠٠ جنيه سنوياً للتصليحات . فبكم جنيها يجب أن يؤجره شهرياً حتى يحصل على ربح صاف قدره ٨٪ من ثمن البيت ؟ (الجواب ————————)

ثانياً : اختبار العمليات الحسابية الصورة م :

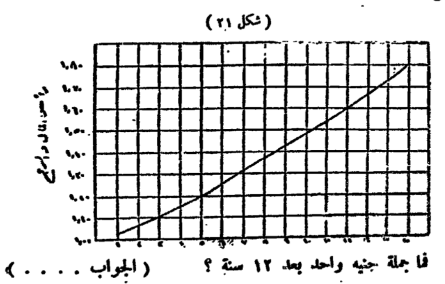
وتشمل ٦٠ فقرة ومثلها الصورة ن .

وإليك مثالا من كل منهما .

الفقرة ٤٩ الصورة (م)

الرسم البيانى الموجود أدناه يبين الجملة فى كل سنة لمبلغ جنيه واحد يربح ربحاً مركباً بسعر ٤٪ فى السنة

(شكل ٢١)



فما جملة جنيه واحد بعد ١٢ سنة ؟ (الجواب . . . .)

الفقرة (٣٨) و(٣٩) الصورة (ن)

أبعاد بعض المدن الشهيرة عن القاهرة بالكيلومتر

(شكل ٢٢)

دمياط
الخرطوم
بورسعيد
الإسكندرية
أسيوط

من هذا الرسم البيانى أوجد بالتقريب بعد المدن الآتية عن القاهرة بالكيلومتر :

دمياط .... الخرطوم ....

العلــوم

نكتفى فى هذا القسم بذكر بعض الأمثلة من أهم الاختبارات الشائعة أو التى كانت شائعة . وأغلب أمثلة الكيمياء والحيوان والنبات والجغرافية مأخوذة عن الأستاذ إسماعيل القبانى فى كتابه مشكلة الامتحانات.

اختبار رتش وبوبينو فى العلوم العامة

وله ثلاثة صور :

« الصورة أ » :

وفيها قسمين نأخذ أمثلتنا من :

القسم الأول :

التعليمات : « ضع خطاً تحت كلمة واحدة أو عبارة واحدة تعتبرها أحسن جواب للسؤال ، كما فى المثال الآتى :

« من أمثلة السوائل : الحديد ــ الزجاج ــ الهواء ــ الماء ــ الرصاص ــ الخارصين ــ الأكسجين » .

ويلى ذلك ٥٠ سؤالا من هذا النوع ، يخصص للإجابة عنها ١٥ دقيقة ، نكتفى بإيراد خمسة منها على سبيل التمثيل ، وهى :

٢ ــ الأقرب إلى الاحتمال أن توجد البراكين فى : الصحارى ــ الوديان ــ السهول الساحلية ــ الجبال ــ دالات الأنهار ــ المناطق الداخلية ــ الجزر .

٧ ــ من أمثلة التأكسد : صدأ الحديد ــ تحليل الماء بالكهرباء ــ انصهار الجليد ــ ذوبان الملح ــ تأثير حامض فى الرخام ــ تسخين كلورات البوتاسيوم ــ غليان الماء .

١٠ ــ عندما يحتوى السائل على أكبر كمية يمكنه إذابتها من مادة ما تسمى هذه الحالة : الأنتشار الغشائى ــ المسامية ــ الانصهار ــ التعادل ــ الاختزال ــ الإفراز ــ التشبع .

القسم الثانى :

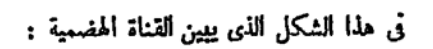
التعليمات : يوجد فى هذه الصفحة والصفحات التالية ٢٠ شكلا بجوار كل منها عبارات متعلقة به ، ولكن قد تركت فى كل من هذه العبارات مسافة خالية ، عليك أن تملأها بحيث تجعل العبارة صحيحة ، وافية ، محدودة المعنى . فى معظم الحالات لا تحتاج إلا إلى كلمة واحدة لملء المسافة . وستعطى لك ٢٥ دقيقة للإجابة عن هذا القسم من الاختبار .

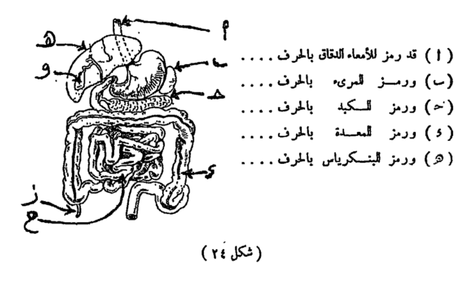
فى هذا الشكل الذى يبين تركيب الزهرة النموذجية :

أ
ب
حـ
ى

(ا) قد رمز للبتلات (التويج) بالحرف . . . .
(ب) ورمز لأعضاء التـذكير بالحرف . . . .
(حـ) ورمز للسبلات (الكأس) بالحرف . . . .
(ى) ورمز لعضو التأنيث بالحرف . . . .

(شكل ٢٣)


فى هذا الشكل الذى يبين القناة الهضمية :



(ا) قد رمز للأمعاء الدقاق بالحرف . . . .
(ب) ورمـز للمرىء بالحرف . . . .
(حـ) ورمز للـكبد بالحرف . . . .
(ى) ورمز للمعـدة بالحرف . . . .
(هـ) ورمز للبنـكرياس بالحرف . . . .

(شكل ٢٤)

اختبار رتش وكرسمان فى علم الحياة

وللإختبار صورتان نأخذ أمثلتنا من الصورة ب . وكل صورة تتكون من ٥ أقسام :

القسم الأول : ٤٠ سؤالا زمنه ١٢ دقيقة . والإجابة بالاختيار من سبعة إجابات عن كل سؤال .

القسم الثانى : ١٥ عبارة زمنها ٨ دقائق . والإجابة بتكميل كل عبارة .

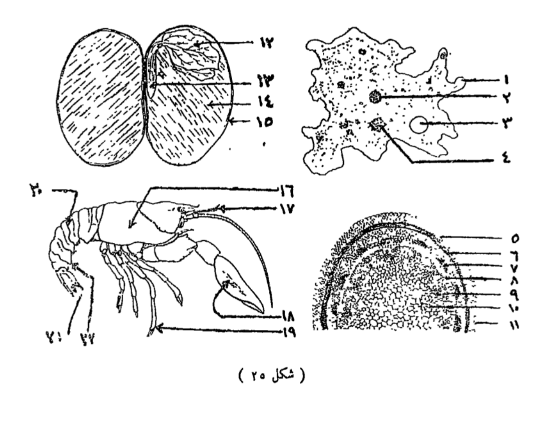
القسم الثالث : التعليمات : افحص الأشكال الأربعة المرسومة على الصفحة المقابلة لهذا جيداً . وقد وضعت على بعض الأجزاء فيها أرقام خاصة ، وفيما يلى أسماء بعض هذه الأجزاء وعليك أن تضع أمام كل اسم الرقم الموضوع عليه فى الرسم الآتى (صفحة ٥٨٢) :

القسم الرابع : سؤال واحد عن قانون مندل فى الوراثة والإجابة بتكميل جدول .

القسم الخامس : التعليمات : اكتب الكلمات المتروكة فى الفقرات التالية ، بحيث تختار لكل مسافة بيضاء الكلمة التى تعطى أصح معنى ، كما فى المثال الآتى :

الشرب من إناء واحد يسبب انتشار . . . الأمراض . . . ويلى ذلك خمس فقرات يخصص لإكمالها عشر دقائق ، ونكتفى بإيراد الفقرة الآتية منها على سبيل التمثيل :

١ - للكشف عن وجود البروتين فى مادة ما تسحق هذه المادة جيداً ، ثم تغلى فى ماء قد أضيفت إليه قطرات من . . . . فإذا ظهر لون. . . . دل ذلك على وجود البروتين . وإذا أضفنا كمية كبيرة من . . . . يتحول هذا اللون إلى . . . .



(شكل ٢٥)

مثال البطن ٢٠ :

النواة . . . .
الفلقة . . . .
الزيلم . . . .
الفجوات الغذائية . . . .
القشرة الخارجية . . . .
اللحاء . . . .
الصدر الرأسي . . . . .
الكامبيوم . . . .

الأرجل القابضة . . . . .
النقير . . . .
القشرة الداخلية . . . . .
الفجوات القابضة . . . .
شعيرات الاستشعار . . . .
الزعانف الذنبية . . . . .
الألياف . . . .

اختبار جلن وولتون فى الكيمياء

التعليمات ثلاث من الكلمات الواردة فى كل سؤال من الأسئلة الآتية بينها علاقة متينة من حيث المعنى أو من حيث الاستعمال أو كليهما . والمطلوب منك أن تكتب بين القوسين الموضوعين فى نهاية كل سؤال رقمى الكلمتين أو الجملتين اللتين لا توجد بينهما وبين تلك الكلمات الثلاث أى علاقة .

ويلى ذلك ٣٠ سؤالا نورد منها الأمثلة الآتية :

(١) ١ — يد ا ص ٢ — يد ا بو ٣ — (يد ا)٢ كا
٤ — ماء النار ٥ — ا٢ ك (.......)

(٢٦) ١ — أزوتات ٢ — يد٣ ز ٣ — ا٢ ك
٤ — ملح شيلى ٥ — الحرير الصخرى (.......)

(١٣) ١ — التحليل الكهربائى ٢ — الكبريت ٣ — الرخام
٤ — السكر ٥ — حامض الكبريتيك (.......)

(٢٧) ١ — آحادى التكافؤ ٢ — أكسجين ٣ — صديوم
٤ — كلسيوم ٥ — كلور (.......)

(٣٠) ١ — الاختزال ٢ — جرام ٣ — أكسيدات
٤ — محلول ٥ — فلزات (.......)

اختبار ويديفليد ووالتر فى الجغرافية

التعليمات : فى العمود الأيمن عبارات تدل على أشياء يعملها الناس فى مختلف الأقاليم ، وفى العمود الأيسر عبارات تدل على بعض أسباب ذلك . فمثلا العبارة رقم ٣ من العمود الأيسر تدل على سبب اشتغال أهالى

النرويج بصـــيد السمك . فعليك أن تكتب رقم (٣) بين القوسين الموضوعين بعد عبارة « صيد السمك فى النرويج » الواردة فى النموذج . اعمل مثل هذا مع باقى العبارات الآتية : —

النتائج		الأسباب
نموذج : صيد السمك فى النرويج .	( )	(١) جبال عالية مكسوة بالثلوج .
(١) حياة التجول عند الأكسيمو.	( )	(٢) أناث يعتمدون فى غذائهم على الحيوانات البرية .
(٢) استعمال الجمال كواسـطة للانتقال .	( )	(٣) أرض مجدبة صخرية مغطاة بالغابات قرب البحر .
(٣) بيوت من الخشب وقشور الأشجار، لها سقوف مائلة من أوراق الشجر .	( )	(٤) جو بارد رطب وكثير من الأراضى المنخفضة الرطبة .
(٤) لبس الناس أحذية خشبية فى أثناء تأدية أعمالهم .	( )	(٥) جو حار جاف وكثير من الرمال الناعمة .
		(٦) جو حار ممطر وكثير من أشجار « الرافيا » .
		(٧) سهول مسطحة وأرض خصبة ومطر معتدل فى الصيف .

مجالات أخرى

أصبحت الاختبارات التحصيلية تطبق فى أغلب الميادين التعليمية . ففى المدرسة الإبتدائية والإعدادية نجد الاختبارات العامة أما فى الثانوية فنجد اختبارات أكثر تخصصاً بينما فى الجامعة تكون اختبارات متقدمة متخصصة

شاملة ويكون بعضها عاماً . وهى هنا أو هناك قد تدخل فى عداد فئات أخرى من الاختبارات سيأتى ذكرها فى القسمين التاليين .

وتعتبر سلسلة الاختبارات التحصيلية المتآذرة Coordinated Series of Achievement Test من السلاسل التى تصلح للاستخدام فى مختلف المناهج الدراسية . ومن أفضلها ثلاث اختبارات تطبق من أجل الأغراض التالية :

١ ــ الاختيار وقبول طلبة الجامعة الجدد .

٢ ــ تقييم المناهج التعليمية .

٣ ــ تقييم الاختبارات المنشورة فى مصادر أخرى والسمة الأساسية المميزة لهذه السلسلة هى استخدامها لمعايير موحدة مقارنة فى كل الاختبارات . وهذا يسهل المقارنة المباشرة بين الدرجات المختلفة المعطاة من مواد دراسية مختلفة إلا أن هذه السلسلة لم تقنن كلها على نفس التلاميذ ، ذلك أن هناك تخصص فى مثل هــذه المستويات التعليمية . فبينما يدرس طالب اللاتينية عامين ، يدرس آخر الأسبانية عامين أيضاً ، ومن هنا لا يمكن إعطاء الطالبين معا اختباراً فى اللاتينية أو فى الأسبانية .

وكانت الدرجات الخام فى كل اختبارات هذه السلاسل الثلاث تحول إلى درجات معيارية متوسطها ٥٠٠ وانحرافها المعيارى ١٠٠ : والجماعة المرجعية التى أخذ المتوسط والانحراف المعيارى من درجات أفرادها تكونت من ١٠٦٥١ تلميذاً طبق عليهم اختبار الاستعداد المدرسى Scholastic Aptitude Test وكانت درجات التلاميذ ، بعد إجراء هذا الاختبار ، التى حصلوا عليها فى السنوات التالية تقارن بدرجات الجماعة المرجعية . ثم يحسب الفرق بين متوسط التلاميذ الجدد ومتوسط الجماعة المرجعية ، ومن دلالة هذا الفرق يقيم المنهج التعليمى الذى شاركوا فيه .

وقد أصبحت هذه السلاسل تقليدا منتظما ، إذ تطبق على الطلبة الجدد المتقدمين لكليات معينة . بل وأصبح بعضها صالحاً للاستخدام فى أغراض التوزيع والتصنيف العام . ويجرى تعديلها بصفة دورية حتى تناسب المناهج الدراسية بالمدارس الثانوية ، وحتى تتفق مع التغيرات التى تطرأ على تعليم المواد الدراسية الآتية :

١ – اللغة الرسمية فى البلاد .

٢ – اللغات الأجنبية .

٣ – العلوم الطبيعية .

٤ – الرياضيات .

٥ – التاريخ .

٦ – المواد الاجتماعية الأخرى .

وقد اختيرت الجماعة المرجعية التى حسب منها المتوسط والانحراف المعيارى بحيث كان أفرادها تلاميذا متوسطى القدرة والتحصيل . ذلك بالاقتصار على من حصلوا على درجات متوسطة فى اختبار ذكاء صادق ، باعتبار هذا محكا لتوسط قدراتهم . كما كان يقتصر على الطلبة من أعمار يغلب أن يكون عليها تلاميذ الفرق الدراسية المتضمنة فى الجماعة ، باعتبار هذا محكا لتوسط تحصيلهم .

كما أخذ محك آخر لتوسط تحصيل جماعات التقنين ، وهو حصول أفرادها على درجات متوسطة فى بطارية تحصيل عام أخرى صادقة . وكذلك اتبع تقليد جديد وهو وضع الفقرات على أساس المدة التى قضاها التلميذ المتوسط فى تعلم المادة المقاسة . فمثلا كانت الجماعة المرجعية فى اختبارات اللغة الإنجليزية قد درست اللغة أربع سنوات بالمدارس الثانوية . بينما كانت الجماعة المرجعية بالنسبة لاختبارات اللغات الأجنبية قد درست هذه اللغات مدة سنتين .

مقاييس الإنتاج

الاتجاه إلى استخدام مقاييس الإنتاج Producuct Scales يعكس لنا تحولا فى طرق تقييم التلميذ . تحول إلى الإجراء ، وقرب إلى واقع الحياة المدرسية المعتادة ، وتأهيل للطلبة حتى يستطيعوا مواجهة الحياة العملية بعد الانتهاء من الدراسة . وأغلب موادها تعتمد على تقييم الخط الكتابى والإنشاء والقصص والوصف وفك وتركيب آلات واستخدام أجهزة وأدوات .

مقياس فان واجنن للإنشاء الإنجليزية

يعتبر مقياس ڤان واجنن للإنشاء والإنجليزية Van Wagnen English Composition Scale مثالا جيداً لهذا النوع من المقاييس . وفيه اختبارات منفصلة للوصف والقصص . وتعطى للطلبة درجات عن مستوى المادة المعروضة ، وبنائها ، والأسلوب الفنى فى الكتابة . ثم تؤخذ متوسطات هذه الدرجات ، وتكون درجة التلميذ الكلية فى المقياس هى متوسط درجاته الثلاث .

وقد أعدّ لويس Lewis مقياساً تحسب درجة الطالب فيه بناءاً على درجاته فى المجالات الخمس التالية :

١ ــ ترتيب الكلمات .

٢ ــ ترتيب ورود ومواضع العبارات .

٣ ــ الجانب الاجتماعى فى القصة .

٤ ــ وصف الظواهر الاجتماعية .

٥ ــ بساطة العرض .

والعيب الأساسى فى أغلب مقاييس الإنتاج أنها لا تراجع حتى تتفق مع تغير محتوى المناهج الدراسية والأحداث القومية والعالمية الهامة . ولكنها وسيلة جيدة يمكن الاعتماد عليها ضمن اختبارات التحصيل المقننة ، باعتبارها جزءاً موضوعياً لتقييم مواد كيفية .

أغراض تعليمية أوسع

ونقصد بها أهداف التربية البعيدة مثل : إنماء التفكير الناقد — وتكوين عادات واتجاهات مرغوبة فى تذوق الفن والأدب — والفهم الذكى للأحداث المعاصرة — وتكامل ما حصله المتعلم من مختلف المناهج الدراسية التى اشترك فيها .

الاختبار التعاونى للشئون المعاصرة

الاختبار التعاونى للشئون المعاصرة Cooperative Contemporary affairs Test يمثل هذا النوع من الاختبارات أصدق تمثيل . وهو يطبق فى ساعة على طلبة الجامعة لقياس الميل والمعلومات فى ميادين ثلاث هى :

١ — الشئون العامة ..

٢ — العلوم والطب .

٣ — الأدب والفن .

ويعدل الاختبار كل سنة حتى يكون عصريا .

الاختبار التعاونى للثقافة العامة :

الاختبار التعاونى للثقافة العامة Cooperative General Culture Test يعطى لطلبة الجامعة لقياس المعلومات العامة فى أهم خمسة ميادين لمناهج الدراسة الجامعية وهى :

١ ـ التاريخ والدراسات الاجتماعية .

٢ ـ الأدب .

٣ ـ العلم .

٤ ـ الفن .

٥ ـ الرياضيات .

ولكل ميدان جزء منفصل من الاختبار يتطلب ٣٠ دقيقة . وقد اختيرت فقرات الاختبار بحيث يستطيع الإجابة عن أكثر من نصفها من نجح فى دراسة المواد المقررة فى الجامعة . وكل اختبارات السلسلة التعاونية تستخدم المعيار المئينى . والعينة ، والإجراء جمعى . وتعطى درجات جزئيـة وأخرى كلية .

وقد قامت هذه السلسلة بنشر اختبارات تناسب تلامذة المدارس الثانوية والجامعة . وهى تقيس الميول والاتجاهات والسمات الاجتماعية . وهى تتضمن إلى جوار هذا اختبارات صممت لقياس المجالات التالية :

١ ـ الاستدلال المنطقى .

٢ ـ تفسير البيانات .

٣ ـ تطبيق المعلومات الأساسية فى العلوم التالية :

(أ) الطبيعة .

(ب) الأحياء .

(حـ) الدراسات الاجتماعية .

(د) الأدب .

فمثلا فى تفسير البيانات تقدم مجموعة منها فى شكل جداول أو خرائط أو رسوم متبوعة بسلاسل من الجمل يقوم المفحوص بوضع علامة صحيح أو خطأ عليها بناءً على مدى كفاية البيانات فى تأكيد ما أتت به هذه الجمل .

أما اختبارات تطبيق المعلومات الأساسية فهى تقتضى استخدام الأسس العلمية فى وصف مواقف الحياة اليومية . وموادها غير موجودة فى الكتب المدرسية : وكل مجموعة من الفقرات متبوعة بخاتمه تصنف حسب رأى الطالب إلى : صحيحة ، تحتمل الصحة والخطأ ، خاطئة . وعلى المفحوص أن يبرر رأيه باختيار أسباب منطقية تدعمه من القطع المكونه من جمل تفسر موقف الحياة اليومية المعروض له .

التفكير الناقد فى الدراسات الاجتماعية

اختبار التفكير الناقد فى الدراسات الاجتماعية Critical thinking in the Social Studies ، قدمه ريتستون Wrightsone . ويستخدم فى المدارس الثانوية والمعاهد . وقد انتقد هذا الاختبار لأن عنوانه مضلل فهو لا يقيس التفكير الناقد فعلا ، بل ولم يقدم بيانات عن صدقه وثباته . والقيمة الأساسية لهذا الاختبار فى أنه يقترح طرقا يمكن استخدامها فى عمل اختبارات مماثلة . وهو يتكون من الأجزاء الثلاث التالية :

١ ــ الحصول على الحقائق .

٢ ــ استخلاص استنتاجات من هذه الحقائق .

٣ ــ تطبيق الحقائق العامة .

والجزء الأول هو اختبار لمهارة مذاكرة الدراسات الاجتماعية باستخدام الخرائط ، والرسوم البيانية ، والقوائم ونقلها من مصادر الحقائق .

والجزء الثانى ، يشبه اختبارات فهم القراءة ، ويتطلب من المفحوص أن يختار ما يمكن أن يستفيد به من كل قطعة على حدة .

وفى الجزء الثالث يعاد تقديم تلك القطع مختلطة بقطع أخرى ويطلب من المفحوص أن يسترجع من كل قطعة من الأولى جملتين ويربطهما معاً

اختبار واطسن ـ جلاسر لتقدير التفكير الناقد

لم يعن اختبار واطسن ـ جلاسر بتقدير التفكير الناقد Watson – Glasser Critical thinking appraisal test فى ميدان بذاته من ميادين المعرفة . وتتضمن آخر صورة معدلة لهذا الاختبار الأجزاء الخمسة الآتية :

١ ـ الاستدلال .

٢ ـ تحقيق الفروض .

٣ ـ الاستقراء .

٤ ـ التفسير .

٥ ـ التقييم .

والجزء الأول تعرض فيه قطع يتبعها استدلال منها ويصنف المفحوص هذا الاستدلال إلى صحيح أو خطأ . والجزء الثانى به فروض وجمل والمطلوب معرفة ما إذا كانت هذه الجمل تؤيد تلك الفروض أم لا ، وما درجة تأكيدها . والجزء الثالث يتطلب القيام بعمليات الاستقراء العقلى . أما الجزء الرابع فيطلب فيه من المفحوص أن يدلل على ما إذا كانت الاستنتاجات المدونة مشتقة من قطعة معينة أم لا . والجزء الخامس والأخير يطلب فيه من المفحوص أن يقرر ما إذا كان بناء القطعة قويا أم ضعيفا .

والمشكلة فى تقدير إجابات المفحوصين هى هل يعد الخطأ خطأ فعلا وكليا ؟ وهل يعد الصحيح صحيحا فعلا وكليا ؟ ولهذا قام واضعو الاختبار بدراسة فقراته والاقتصار على أصدقها وأكثرها صلاحية . وكانت طريقة الدراسة بعرض الفقرات على مجموعة من المحكمين عددهم ٣٥ قاض مختارين ممن نالوا تدريبا عالياً وتفوقوا فى المنطق ، وفقه اللغة ، والكيمياء ، والأحياء ، والطبيعة ، وعلم النفس ، والتربية ، والهندسة التطبيقية . فإذا

نالت الفقرة أغلبية أبقى عليها . وقد تم اختبار الفقرات على أساس آخر أيضاً هو الاتساق الداخلى . وفى بحث الصدق استخدمت طريقة الجماعات المتقابلة Contrasted groups وهى جماعة من الطلبة الجامعيين فى مقابل جماعة من العمال والموظفين .

والمعيار المستخدم هو الدرجات المئينية . ومعاملات الثبات منخفضة بالنسبة لبعض الدرجات الفرعية . أما ثبات الاختبار ككل فيتراوح بين ٨٠ر٠ إلى ٩٠ر٠ .

الاختبارات التحصيلية المهنية

يجب ألا ننس أن الكثير من الاختبارات التحصيلية كان يستخدم فى الاختيار ، والتصنيف فى الصناعة ، والحكومة ، والقوات المسلحة . واختبارات التحصيل المهنى Vocational Achievement Tests فى ميدان الصناعة تسمى اختبارات الحرف trade tests . وكثير من الاختبارات التحصلية المهنية وضعت لتؤدى أغراض خاصة ، أى أنها « تفصيل tailor made » . وتحوى موادا تختلف فى طبيعتها ، فقد يكون المحتوى كله لفظيا ، وقد يكون رسومات بيانية أو صوراً ، أو أسئلة تقدم شفويا أو كتابيا .

وهى تستخدم عينات تقنينية من العمال ، ويطلب من المفحوص القيام بعمل يشبه العمل المعتاد الذى يؤديه . وهذا يعنى أن الفقرات عينات من العمل Work Samples . وقرب السلوك المقاس من العمل اليومى ، وتمثيل عينة الفقرات لوحدات السلوك المراد قياسه ، أمران مميزان لهذا النوع من الاختبارات .

ويسجل أداء المفحوص فى اختبارات عينة العمل بطريقة موضوعية . فمثلا ، فى اختبار الآلة الكاتبة ، يحسب مجموع الأخطاء فى النسخة المكتوبة ،

والزمن ، والتنسيق . ولكن بعض هذه الاختبارات يلجأ ، فضلا عن التصحيح الكمى ، إلى التفسير الكيفى الذى يقوم به الخبراء .

ومن أفضل الاختبارات التحصيلية المهنية تلك الطبقة فى المهن المعروفة وخاصة الآلة الكاتبة ، والعنونة ، والسكرتارية . ويتضمن بعض هذه الاختبارات استخدام اللغة الرسمية ، واللغات الأجنبية ، والمعلومات العامة عن المهن ، ومقاييس للمهارة فى المهنة .

اختبار بلاكستون للكفاية فى السكرتارية

اختبار بلاكستون للكفاية فى الأعمال المكتبية Blackstone Stendeographic proficiency Test : يطلب من المفحوص تسجيل ملاحظات مختزلة هى تلخيص واختزال وترجمة لخطابين . وهناك جزء لقياس الأخطاء فى القواعد ، والآخر فى الهجاية ، وثالث فى الضبط ، ورابع فى الربط ، وخامس فى التدريب على الأعمال المكتبية وتنظيم المكاتب .

اختبار سيشور — بينيت للكفاية فى السكرتارية .

يتطلب اختبار سيشور — بينيت للكفاية فى السكرتارية Seashore — Bennett Steneogrephic Proficiency Test — أن يدون المفحوص ملاحظات مختزلة عن ٥ خطابات تتدرج فى الطول والتعقيد . ويكون إملاء الخطابات بسرعات تتدرج فى الزيادة أيضاً . ويطلب من المفحوص أن يسجل ملاحظاته هذه بطريقة تسهل تدوينها بالآلة الكاتبة . والمعايير غير واضحة ، ولكن الصدق مرتفع باتباع طريقة الجماعات المتقابلة ، وبإيجاد معاملات الارتباط بين الدرجات على الاختبار وتقديرات المشرفين على العمل لنفس الأفراد المختبرين ؛ والثبات مرتفع بالصور المتكافئة .

اختبارات الحرف

اختبارات الحرف الشفوية Oral trade Tests هى اختبارات تعين فى القيام بالمقابلات فى مكاتب العمل . وقدا استخدمت هذه الاختبارات على نطاق واسع فى الحربين العالميتين : وقد وصف تشابمان Chapman أول اختبار حرفة شفوى فى الحرب العالمية الأولى ، فى كتاب بأكمله . وقام طومبسون Thompson بمراجعة فقرات هذا الاختبار ووضعها فى قالب جديد .

كما قام قسم تحليل العمل بمكتب العمل بالولايات المتحدة الأمريكية ، حديثاً ، بتأليف اختبار مماثل مقنن يصلح لقياس الكفاية فى ١٢٦ عملا . ولكل عمل ١٥ سؤالا . ووضعت صور مكافئة بالنسبة لبعض هذه الأعمال وعددها ٤٩ عملا . كما حسبت معاملات الثبات بطريقة الصور المتكافئة وتراوحت بين ٠,٧٩ إلى ٠,٩٣ .

وجمعت البيانات عن صدق هذه الاختبارات بملاحظة الأعمال والاختلاط بالعمال و الملاحظين والفنيين فى كل عمل . والأسئلة موضوعة بلغة سهلة ومبسطة بحيث تناسب مستوى العمال . وإجابة بعض الأسئلة تقتضى السرعة ، ويقتضى البعض الآخر إصدار أحكام نافذة . وقد أدى الاختلاط بالعمال إلى مراجعة فقرات الاختبارات وتعديل بعضها وإضافة البعض الآخر . وروعيت الفروق بين المناطق فى مواد العمل وطرق أدائه .

وقد قسم العمال إلى هذه الأقسام :

(ا) الأسطوات : الذين قال رؤساؤهم إنهم أعلى مستوى وهم أصحاب خبرة ٤ سنوات على الأقل فى عمل بذاته :

(ب) المبتدئون : كالصبيان المساعدين .

(جـ) من يعملون فى أعمال متشابهة ويقرب مستواهم من الأسطوات .

وقد قننت الأسئلة كل سؤال على عمل من ٥٠ إلى ١٠٠ عامل من الجماعة ا ومن ٢٥ إلى ٥٠ من الجماعتين ب ، ج .

وقد اختيرت الأسئلة على أساس القيمة التقريبية .

ونظراً للتغير الدائم فى العمليات والمواد المهنية فمن الواجب دائماً أن تراجع اختبارات الحرف ولكنها على أى حال لا تقوم مقام الأداء فى اختبارات عينة العمل . ولكنها تستخدم وحسب كمقياس سريع لمعرفة الخبرة المهنية للباحث عن عمل .

الفصل الثامن عشر

بحوث واختبارات تربوية محلية

- مقـــدمة .
- بحث الأستاذ للقبانى .
- بحوث وزارة التربية والتعليم .
- بحوث المؤلف .
- دليل الاختبارات .
- القراءة .
- الحساب .
- نحو المجلد الثانى .

الفصل الثامن عشر

بحوث واختبارات تربوية محلية

مقـــدمة :

قد يعتقد البعض أننا لن نلحق ركب العلوم الموضوعية عامة والقياس خاصة ، ولكننا نعتقد أن هذا البعض عندما ينتهى من قراءة هذا ا ل سيكون أكثر تفاؤلا بمستقبل البحوث والدراسات فى بلدنا . ونحن لا نهدف إلى تدعيم هذا التشاؤم أو ذاك التفاؤل ، بل وإننا نحاربه ، إذ أن اعتقاداً متطرفاً من أحد النوعين لن يقدم لنا جديداً فضلا عن تثبيطه همم البحاثة وعدم إعطاء الحق لأصحابه .

وإن كنا نثق بأنفسنا ونفخر بجهود روادنا ، إلا أننا لم نقصد باختيار وتقديم بعضها تقييما أو تقـــديراً ، بل تمثيلا يعطى فكرة للقارئ ويستثير الكفاءات لتنهج نهجه . على أن البحوث التربوية التى استخدمت الاختبارات أداة لها طرقت مجالات لها فى بلادنا تاريخ طويل ، وإن كانت قد بدأت متعثرة بعض الشيء إلا أن هذا هو شأن الجهود العلمية فى ميدان جديد نسبياً كما أن إمكانيات هذه البحوث ، مهما كانت الهيآت القائمة بها ، قتيرة إذا ما قورنت بمثيلاتها فى الخارج . إلا أنها تخطو خطوات واسعة وتتقدم تقدماً سريعاً .

ونأمل أن يحين الوقت الذى فيه تكون ثروتنا من البحوث والدراسات فى المجال الذى يمر به كل فرد ويعيش فيـــه الكثير من العاملين وهو ميدان التربية .

بحث الأستاذ إسماعيل القبانى

قدم الأستاذ إسماعيل القبانى بعض أسئلة امتحانات الشهادات العامة حتى سنة ١٩٣٨ . ثم عقب عليها مقدماً بحثه وملخصاً إياه فى كتاب مشكلة الامتحانات فى مصر التى أصدرته رابطة التربية الحديثة ١٩٣٩ قائلا :

إن أمثال هذه الأسئلة يجدها الإنسان فى كل ورقة من أوراق أسئلة المواد التى تعتبر « مواد معلومات » . وظاهر أن الإجابة عنها لا تتطلب من التلميذ أكثر من أن يحفظ بضع صفحات من الكتاب المقرر عليه ، ويستبقيها فى ذهنه إلى أن يحين وقت الامتحان ، فيفرغها على ورقة الإجابة كما هى من غير تصرف . وهذا يجعله أقرب إلى قياس الملكة اللفظية وذاكرة الألفاظ .

ولا يقتصر الأمر على هذا ، بل هناك ما هو أهم وأدهى : فإن مثل هذا الامتحان لا يمكن أن يكون مقياساً صحيحاً للتحصيل وللقدرات العقلية المكتسبة من التعليم ، مع أن هذا هو الغرض المباشر من إجرائه ؛ فما كان المقصود بالتحصيل العلمى يوماً من الأيام شحن الذاكرة بحقائق غير مهضومة وغير مترابطة ، تدخل الذهن من ناحية ثم تخرج من الناحية الأخرى ولم تغذ فيه قوة نافعة .

إن أكبر ما يؤخذ على التعليم فى مدارسنا هو أنه لا يثقف عقل الطالب ولا ينمى فكره ، ولا يربى فيه قوة الابتكار وسلامة الحكم ، وروح النقد ، والاستقلال فى الرأى ؛ ولا يبعث فى نفسه حب العلم ، وتقديس الحق ، والشعور بالواجب ، وبالإجمال لا يهيئ الفرص لتكوين الشخصية ، وبناء الخلق المتين ، الذى يؤهل الطالب للكفاح فى الحياة ، ولخدمة المجتمع . ومن أهم أسباب ذلك أننا أخطأنا غاية التربية ، فصرنا نعنى بتكديس المعلومات ، ولا نبحث على أثرها فى النفس . ولا شك فى أن

الامتحانات بصورتها الحالية عامل قوى ، بل هى أقوى عامل ، يوجه التعليم فى ذلك الاتجاه .

إن الصورة التى يكونها المرء فى ذهنه عن الحياة العقلية للطالب المصرى عند ما يفحص أوراق الأسئلة يتمثل فيها العقل فى شكل صندوق مستقبل تصب فيه المعلومات كما يصب الماء فى مستودعه ، وتبقى فيه إلى أن يحين وقت تفريغها منه كما يسحب الماء من المستودع . ولكن المـاء يبقى فى المستودع مستقلا عنه ، لا يؤثر فى تكوينه ، فهل هذه هى العلاقة التى نريدها بين المعلومات والعقل ؟ أم نريدها علاقة الغذاء بالجسم ينميه ويقويه ؟ والغذاء لا يبقى فى الجسم على حالته الأصلية ، بل يتفاعل مع إفرازاته ويتحول إلى عصارة مهضومة ، تختلف فى تركيبها عن الطعام الأصلى ، ويمتصها الجسم فيكسب منها القوة ، أما إذا بقى الطعام فى المعدة على حالته الأولى فلا تنتج عنه إلا تخمة . فكذلك المعلومات إذا بقيت فى الذهن عائمة منفصلة ، مستعدة للخروج منه على نفس الصورة التى دخلت بها ، فإنه لا تنتج عنها إلا تخمة عقلية . فالمعلومات التى يحصلها الإنسان لا تصبح جزءاً من نسيجه ، فينمو بما اكتسبه منها ، ويصير بعد امتصاصها أقدر على التصرف فى المشاكل التى تعرض له مما كان قبلا ، وبذلك تؤثر فى اتجاهات الإنسان النفسية ، وتصبغ نظرته العامة إلى أمور الحياة من غير أن يشعر . وفى هذه الحالة لا يصبح من المتيسر اختبار قيمة هذه المعلومات باستعادتها من العقل كما كانت ، بل بقياس القوة التى اكتسبها العقل منها .

وقد أحسن التعبير عن هذا المعنى الفيلسوف الإنجليزى الرياضى الأستاذ هوايتهد (Whitehad) عند ما حذر المربين فى كتاب ممتع له عن التربية مما سماه « الأفكار الخامدة » (inert idees) وقد عرفها بأنها « الأفكار التى يقتصر العقل على استخدامها ، من غير أن يستخدمها ، أو يختبر صحتها ، أو يحولها إلى صور جديدة » . ثم فسر ما يقصده باستخدام الأفكار بقوله :

« أقصد باستخدام الفكرة إيجاد الصلة بينها وبين ذلك التيار الذى يتألف من إحساسات ووجدانات وآمال ورغبات ومن نشاط عقلى يلائم بين نواحى التفكير المختلفة ، وهو الذى تتكون منه حياة الإنسان » .

إن أهم انقلاب فى أساليب التربية الحديثة ، وأهم غرض يقصد من الطرق التى ابتكرها النابهون من المربين ، كطريقة المشروعات ، وطريقة دكرولى ، وغيرهما ، هو العمل على تحرير التعليم من عبادة الأفكار الخامدة ، وتحويله إلى التثقيف العقلى وتنمية الشخصية . ولكن الامتحانات التى هى من قبيل امتحاناتنا عقبة كؤود فى سبيل تطبيق هذه الأساليب ، لأنه ما دام الامتحان يتطلب سرد حقائق مستظهرة من غير تصرف ، ومادام المعلمون والطلاب وأولياء أمورهم والمشرفون على التعليم بوزارة المعارف (فى هذا الوقت) يحرصون — وطبيعى أن يحرصوا — على نجاح الطلاب فيها ، فالمدارس لا تتحرج من أن تصرف الوقت فى حشو أذهانهم بالأفكار الخامدة وتضن بهذا الوقت أن يصرف فى مطالعة حرة ، أو بحث شخصى ، أو نشاط اجتماعى ، أو هواية عملية ، أو رياضة بدنية ، أو غير ذلك من نواحى التربية الصحيحة ، التى تنمى الشخصية وتوسع الفكر . ولو أن أسئلة الامتحان كانت تتطلب إعمال الفكر ، وسلامة الحكم ، وحسن التصرف فى استخدام المعلومات ، لاتجه التعليم إلى العناية بهذه الصفات ووجوه النشاط التى تنميها . فأهم ما تمس إليه الحاجة إذن من وجوه الإصلاح فى الامتحانات عندنا هو إصلاح نوع الأسئلة .

وليس أثر الامتحانات السيئ فى توجيه التعليم مقصوراً على ما سبق ، بل إن الامتحانات العمومية من أقوى عوامل الجمود فيه . لأنها تتطلب توحيد المناهج والأساليب ، فتحول دون مراعاة المدراس لحاجات ... تمشيا مع الحوادث التى تثير اهتمام المجتمع ، وتمنعها من مسايرة

ميول تلاميذها وإطلاق الحرية المعقولة لهم فى الدراسة . كما أنها تحد من حرية المعلمين فى اختيار مادة الدروس وطرق التدريس ، لأن المدرس لا يعرف مزاج الممتحن الخارجى ، وطريقته فى وضع الأسئلة ، ووجهة نظره فيما يتعلق بالأهمية النسبية لموضوعات المنهج ، والمدى الذى ينبغى الذهاب إليه فى تدريس كل منها ، ولذلك يخشى أن يعتمد على حكمه الشخصى فى العناية بنقط معينة يراها هامة ، أو إغفال نقط أخرى يراها ثانوية أو غير ملائمة لمستوى تفكير تلاميذه وميولهم ، لئلا تجىء أسئلة الامتحان فيما أغفله . ونتيجة ذلك أن أكثر المدرسين يعنون بدرجة واحدة بكل كبيرة وصغيرة ، ويصرفون الوقت فى تدريب تلاميذهم على أنواع معقدة من الأسئلة والتمارين لا يؤمنون بفائدتها ، ولا يقصدون منها إلا سد الأبواب على الممتحن الخارجى .

ثم إن شبح الامتحانات الماثل دائما للأذهان يدفع المدرسين إلى إرهاق التلاميذ بالعمل ، والإكثار من التكاليف والواجبات المنزلية ، ويدفع الكثيرين من الآباء إلى الضغط على أبنائهم ضغطا مستمرا لحملهم على الاستذكار . غافلين عما يلزم لصحتهم من راحة ورياضة ، ويدفع التلاميذ عند اقتراب موعد الامتحان إلى وصل الليل بالنهار فى مجهود مضن ، ينتهى فى كثير من الأحيان بإضعاف أجسامهم وإتلاف أعصابهم .

وهناك ناحية أخرى لموضوع الامتحانات . وهى البحث عن مبلغ دقتها من حيث هى أداة للقياس . وبعبارة أخرى البحث عما إذا كانت الدرجة التى يحصل عليها طالب فى مادة من المواد تدل بالضبط على حقيقة قوته فى تلك المادة . إن أقل خبرة بالامتحانات المألوفة لتشعر بأن درجاتها بعيدة عن الدقة فى قياس قوى التلاميذ ، ولكن مدى الأخطاء التى تتعرض لها هذه الدرجات لم يظهر بجلاء إلا نتيجة للأبحاث التجريبية التى قام بها العلماء . وقد أجرى معهد التربية بإشراف الأستاذ القبانى فى شهرى يناير

وفبراير سنة ١٩٣٠ تجربة من هذا القبيل على امتحانات المدارس المصرية ، قد يكون من الطريف أن تعرض نتائجها ، لنلمس مقدار الخطأ الذى نقع فيه بتعليق أهمية كبيرة على فروق يسيرة فى الدرجات بين تلميذين أو بين امتحانين للتلميذ الواحد .

والفكرة التى قامت عليها التجربة هى أن من علامات أداة القياس المضبوطة أننا إذا قسنا بها الشيء الواحد مرتين فلا بد أن تتفق نتائج القياسين ولو تغير الشخص الذى يقوم بالقياس ، أو الظروف التى أجرى فيها . فإذا قسنا طول مستقيم مرتين بمسطرة واحدة أو بمسطرتين متشابهتين . فلا يختلف طوله فى المرة الثانية عنه فى المرة الأولى ما دام الشخصان اللذان قاما بالقياس فى المرتين يعرفان كيفية استعمال المسطرة . فبالمثل إذا كان الامتحان مقياسا مضبوطا لما يحاول أن يقيسه ، فإننا إذا أعطينا فصلا من التلاميذ امتحانين متماثلين فى مادة واحدة فى زمنين متقاربين ، وجب أن تتفق نتائجهما . فلتحقيق ذلك انتهز الأستاذ القبانى فرصة امتحان وسط السنة فى شهر يناير سنة ١٩٣٠ ، فاتفق مع نظار بعض مدارس القاهرة الابتدائية والثانوية على إعادة الامتحان لبعض الفصول فى مواد معينة .

وقد أعطى الامتحان الثانى فى الأسبوع الأول أو الثانى عقب العودة من أجازة وسط السنة ، واتخذت كل الاحتياطات الممكنة لضبطه وجعله مكافئا لامتحان وسط السنة ومماثلا له فى جميع الوجوة . فاختير لتأليفه وتصحيحه أقدر مدرسى المادة المختارة فى كل مدرسة . وأعطيت لهم نسخة من امتحان وسط السنة فى تلك المادة وطلب منهم أن يضعوا امتحانهم على نمطه ، وأن يجعلوه مماثلا له فى الصعوبة . وألا تخرج الأسئلة عن الموضوعات التى امتُحن فيها التلاميذ فى وسط السنة ، بحيث أن هذه الأسئلة كان يصح أن تكون هى التى أعطيت فى امتحان وسط السنة لو أن المدرس الذى وضعها كان هو الذى اختير لوضع أسئلته .

وقد أخطر التلاميذ عن الامتحان مقدما فى بعض الفصول قبل إجرائه بمدة تسمح لهم بالمذاكرة والاستعداد له ، وأعطى للبعض الآخر من غير إخطار سابق ، للمقارنة بين النتائج فى الحالتين . وبعد إجراء الامتحان صححت الإجابات بكل عناية ودقة ، ثم قورن بين نتائجه ونتائج امتحان وسط السنة فى كل فصل . وفى الصفحة التالية جدول يحتوى على نتيجة المقارنة ، وقد وضعت العلامة * فوق المواد التى أخطر التلاميذ عن امتحانها مقدما .

وبالنظر إلى هذا الجدول نرى أن المقارنة قد أجريت بثلاث طرق مختلفة وهى :

أولا : حسب عدد التلاميذ الذين نجحوا فى الامتحانين معاً ، وعدد الذين رسبوا فى الامتحانين معاً ، وعدد الذين نجحوا فى أحد الامتحانين ورسبوا فى الآخر . وقد ظهر أن ٤٠ تلميذا من جملة تلاميذ المدارس الثانوية الذين أعيد امتحانهم — وعددهم ١٥٦ — قد نجحوا فى أحد الامتحانين ورسبوا فى الآخر ، أى أن نجاحهم أو رسوبهم فى أى من الامتحانين مطعون فيه . ومثلهم ١٤ من تلاميذ مدرسة العباسية الابتدائية ، وعددهم ٦٥ .

ثانياً : قسمت النهاية العظمى للدرجات فى كل مادة إلى خمس أقسام متساوية وبحث عن القسم الذى تقع فيه درجة كل طالب فى كل من الامتحانين . ومن ذلك أمكن معرفة ما إذا كان الطالب فى الامتحان الثانى قد بقى فى القسم الذى كان فيه فى الأول أو تقدم أو تأخر ، وما عدد الأقسام — أو الخطوات — التى تقدمها أو تأخرها . فمثلا إذا كان الطالب قد حصل فى امتحان وسط السنة فى مادة ما على ٧ درجات من ٢٠ فدرجته تقع فى القسم الثانى ، وإذا كان قد حصل فى نفس المادة على ١٣ من ٢٠ فى الامتحان الثانى فدرجته هذه تقع فى القسم الرابع ، وبذلك يكون قد تقدم خطوتين .

الموازنة بين نتائج امتحانين متكافئين فى مادة معينة للفصل الواحد

المادة	المدرسة	الفرقة	العدد الكلى للتلاميذ الممتحنين	عدد من نجحوا فى الامتحانين	عدد من رسبوا فى الامتحانين	عدد من نجحوا فى أحد الامتحانين ورسبوا فى الآخر	عدد من تقدمت درجاتهم أو تأخرت بقدر : خطوة واحدة	خطوتين	ثلاث خطوات	أربع خطوات	معامل الارتباط بين نتائج الامتحانين
مـدارس ثانويـة											
اللغة الإنجليزية	التوفيقية الثانوية	الرابعة – علمى	٣٣	٢٦	٢	٥	١٤	٣	١	—	٠ر٦٠
الهندسة*	شبرا الثانوية	الثالثة	٢٥	٤	١٤	٧	٦	٣	—	—	٠ر٥٠
الهندسة	التوفيقية الثانوية	الثالثة	٣٢	٦	١٢	١٤	١٤	٥	٣	—	٠ر٥٢
الطبيعة والكيمياء*	شبرا الثانوية	الثالثة	٢٨	٢٠	٥	٣	١٠	٣	—	—	٠ر٦٨
الجغرافية	الإبراهيمية	الأولى	٣٨	١٣	١٤	١١	٢٢	٣	٢	١	٠ر٣٩
الجملة			١٥٦	٦٩	٤٧	٤٠	٦٦	١٧	٥	١	
مدرسـة ابتدائيـة											
اللغة العربية*	العباسية الابتدائية	الثانية	٣٦	٢٢	٦	٨	١٧	٣	--	—	٠ر٦٨
الجغرافية	العباسية الثانوية	الثالثة	٢٩	٢٠	٣	٦	١٣	٢	—	-	٠ر٧٥
الجملة			٦٥	٤٢	٩	١٤	٣٠	٥	—	-	

* فى هذه الحالات أخطر التلاميذ عن الامتحان الثانى مقدما للاستعداد

(جدول ١٣)

ونرى من الجدول أن ٨٩ من تلاميذ المدراس الثانوية الذين أعيد امتحانهم – أى أكثر من نصفهم – قد انتقلت درجاتهم فى الامتحان الثانى إلى قسم غير الذى كانت فيه فى وسط السنة ، ومنهم ٦ انتقلوا ثلاث خطوات أو أكثر ، أى أن درجتهم كانت فى أحد الامتحانين أقل من ٢٠٪ من النهاية العظمى وفى الامتحان الثانى أكثر من ٦٠٪ منها ، أوكانت فى أحد الامتحانين أقل من ٤٠٪ من النهاية العظمى وفى الامتحان الثانى أكبر من ٨٠٪ منها !

وفى مدرسة العباسية الابتدائية نجد أن ٣٥ تلميذاً من ٦٥ انتقلت درجاتهم إلى قسم غـــير الذى كانت فيه ، ومنهم ٥ انتقلت درجاتهم خطوتين .

ثالثاً : حسب « معامل الارتباط » بين درجات التلاميذ فى الامتحانين فى كل مادة .

وظاهر من الخانة الأخيرة فى الجدول أن الاتفاق بين درجات الامتحانين فى مجموعها ليس كبيراً ، بل هو ضعيف جداً فى بعض المواد ، إذ المقرر فى القياس العقلى أنه لا يصح التعويل على نتائج أى اختبار إذا قل معامل الارتباط بين الدرجات التى نحصل عليها من صورتين متكافئتين منه عن نحو ٧٥ ر ٠

ولا يجوز أن ننسب اختلاف الدرجات فى الامتحانين إلى عامل ثابت ، مثل قلة اهتمام التلاميذ بالامتحان الثانى أوكثرة استعدادهم لامتحان وسط السنة . لأن هذا لوصح لترتب عليه نقص غام فى درجات سائر التلاميذ فى الامتحان الثانى عنها فى امتحان وسط السنة ، والواقع غير ذلك ، كما يتبين من النظر إلى الجدول التالى ، وفيه الدرجات التى حصل عليها التلاميذ فى كل من الامتحانين ، فى الحالات التى كان فيها اختلاف كبير . فمن هذا الجدول ترى أن بعض التلاميذ زادت درجاتهم فى الامتحان الثانى عليها فى

التلاميذ الذين تقدمت أو تأخرت درجاتهم فى الامتحان الثانى عنها فى الامتحان الأول بقدر خطوتين أو أكثر

« الخطوة تعادل ٢٠٪ من النهاية العظمى لكل مادة »

المـادة	المدرسة	المفحوص	الدرجة فى امتحان وسط السنة	الدرجة فى الامتحان الثانى
اللغة الإنجليزية النهاية العظمى ٣٠	التوفيقية الثانوية	١	١٠	٢٠
		٢	١٦½	٢٤
		٣	١٧½	٢٤
		٤	١٨½	٥½
الهندسة النهاية العظمى ٢٠	شبرا الثانوية	٥	٦	١٣
		٦	١٦	٩
		٧	٤	١٣
الهندسة النهاية العظمى ٢٠	التوفيقية الثانوية	٨	٧	١٤
		٩	٩	٢٠
		١٠	١٣	٧
		١١	٦	٤
		١٢	١٤	٧
		١٣	٢٠	٧
		١٤	١٩½	٧
		١٥	٢٠	٦
الطبيعة والكيمياء النهاية العظمى ٤٠	شبرا الثانوية	١٦	١٠½	٢٤
		١٧	١٤½	٢٤½
		١٨	١٤	٢٦
الجغرافية النهاية العظمى ١٥	الابراهيمية الثانوية	١٩	٦	٢½
		٢٠	٥½	٩
		٢١	٢½	٦
		٢٢	١½	١٢
		٢٣	٤	١١
اللغة العربية النهاية العظمى ٤٠	العباسية الابتدائية	٢٤	٣٤	٢٠
		٢٥	٣٥	٢٠
		٢٦	١٦	٦
الجغرافية النهاية العظمى ١٥	العباسية الابتدائية	٢٧	٩	٤
		٢٨	١٢	٨

(جدول ١٤)

امتحان وسط السنة ، والبعض الآخر نقصت درجاتهم فيه ، مما يدل على أن الاختلاف يرجع إلى أسباب تتصل بطبيعة الامتحان .

وأهم هذه الأسباب ما يأتى :

أولا : الامتحان العادى فى أية مادة لا يقيس شيئاً واحدا ، بل أشياء كثيرة مختلطة بعضها ببعض ، ومن القواعد المقررة عند علماء القياس العقلى أن المقياس الذى يستعمل لقياس شيئين فى آن واحد لا يقيس شيئاً ما . فامتحان الجغرافية مثلا يختبر تحصيل التلاميذ فى بعض موضوعات هذه المادة ، ويختبر فى الوقت نفسه قوة تفكيرهم ، وقدرتهم على التعبير عن أفكارهم ، وقوتهم فى رسم الخرائط ، وسرعتهم فى قراءة الأسئلة وفى كتابة الإجابة عنها ، غير ذلك من العوامل التى تتوقف عليها الدرجة التى تعطى للتلميذ . وهذه العوامل تختلط بعضها ببعض بنسب تختلف من امتحان لامتحان ، مما يؤدى إلى اختلاف نتيجتها مجتمعة .

ثانياً : ورقة الامتحان العادى تحتوى على عدد قليل من الأسئلة ، لا يمكن أن يتناول جميع موضوعات المنهج الذى يمتحن فيه التلاميذ ونقطه . ولذلك قد يصادف الحظ تلميذا فى أحد الامتحانين فتأتى الأسئلة مما يتقنه ، وقد يعاكسه الحظ فى الامتحان الثانى فتأتى بعض الأسئلة مما لا يعرفه ، ويتسبب من ذلك اختلاف درجته فى الامتحانين . فالدرجة لا تمثل قوة التلميذ فى المادة بأجمعها ، بل فى ذلك الجزء منها الذى اشتملت عليه ورقة الأسئلة .

ثالثاً : الامتحان مقياس ذاتى ، ومعنى هذا أن النتائج التى يعطيها تتأثر برأى الشخص الذى يقوم به وتقديرة . وتتجلى هذه الذاتية فى جميع خطوات الامتحان : فالحكم على درجة صعوبة الأسئلة ومقدار ملاءمتها لقوة التلاميذ هو حكم ذاتى يتوقف على رأى واضعها الشخصى الذى

لا ضابط له . وتقدير الدرجة التى تستحقها إجابة كل تلميذ هو تقدير ذاتى يتوقف على رأى المصحح ، وقد أثبتت التجارب العلمية الكثيرة اختلاف المصححين اختلافا بيِّنا فى تقدير الدرجة للإجابة الواحدة ، بل واختلاف المصحح الواحد في تقدير الدرجة التى تعطى للإجابة إذا عرضت عليه مرات تفصلها فترات طويلة أو قصيرة من الزمن . ثم إن تحديد مستوى النجاح أيضاً هو تحديد ذاتى يتوقف على رأى المشرع ، وهو رأى اعتباطى لا يستند إلى سند معقول ، ولذا نجد الحد الأدنى للنجاح فى بعض الأحيان ٤٠٪ ، وفى بعضها ٥٠٪ ، وفى بعضها ٦٠٪ ، وكل هذه النسب المئوية ليس لها فى الحقيقة مدلول ثابت ، فمدلولها يتوقف على ظروف كثيرة ، من أظهرها مقدار صعوبة الامتحان أو سهولته :

وقد انتهى الأستاذ القبانى من هذا البحث إلى أن نظام الامتحانات بما فيه من عيوب ، كشف عن بعضها البحث ، لا تقل شأنا عن الاختبارات الموضوعية المقننة التى يمكن أن تفيد فى نفس الأغراض . ذلك أن الاعتماد على الامتحانات وحدها لم يعد يفى الغرض ولا كذلك الاعتماد على الاختبارات وحدها .

بحوث وزارة التربية والتعليم

قامت وزارة التربية والتعليم بالإقليم الجنوبى بعدة بحوث ودراسات كما أنشأت الكثير من الاختبارات أدوات لهذه البحوث ، والبحوث ما زالت جارية ، وأخرى يعتزم القيام بها . وقد اشترك المؤلف فى عدد من هذه البحوث والاختبارات وهذا سرد سريع لأهم هذه البحوث :

(ا) البحوث التى قاربت الانتهاء :

١ ـ بحث التقدير الرقمى للموظفين .

٢ ـ بحث تشخيص الأخطاء الإملائية .

٣ — بحث تيسير الكتابة والطباعة فى اللغة العربية .

(ب) البحوث الجارية :

١ — بحث كيفية الانتفاع بوقت الفراغ .

٢ — بحث مستويات الدراسة فى العلوم والرياضة فى المرحلة الإعدادية والثانوية .

٣ — بحث المستوى التحصيلى والتأخر الدراسى فى المرحلة الابتدائية .

٤ — بحث مشكلة نظام القبول بالجامعات .

٥ — بحث تطوير مناهج التعليم فى الأزهر .

٦ — بحث الطلبة المتفوقين فى المدارس الثانوية .

٧ — بحث مقارنة ما ينتج من تعليم لغة أجنبية واحدة أو لغتين .

٨ — بحث النهوض بالتعليم فى مناطق الحدود .

(حـ) بحوث لا زالت فى المراحل الأولى للتنفيذ :

١ — بحث مدى ملاءمة التعليم الابتدائى للبيئة .

٢ — بحث تقويم التعليم الإعدادى الفنى .

٣ — بحث دور المدرسة فى تكوين الوعى القومى وتنميته .

٤ — بحث تعليم الخط العربى فى المدارس الابتدائية .

(د) بحوث لا زالت فى مرحلة التخطيط :

١ — برنامج وحدة بحوث المعاهد الزراعية العليا .

٢ — برنامج وحدة البحوث التربوية بكلية المعلمين .

٣ — برنامج وحدة بحوث العلوم والرياضة بكلية المعلمين .

٤ — برنامج وحدة البحوث الاحصائية .

وهذا بيان بالاختبارات التى أعدتها الوزارة ، والتى تهمنا فى هذا القسم :

اختبارات المستوى التحصيلي للمرحلة الابتدائية فى اللغة العربية :

اختبار القدرة على الفهم فى القراءة الصامتة للصف الثالث .

(صورتان ا ، ب) .

اختبار القدرة على الفهم فى القراءة الصامتة للصف الرابع .

(صورتان ا ، ب) .

اختبار القدرة على الفهم فى القراءة الصامتة للصف الخامس .

(صورتان ا ، ب) .

اختبار القدرة على الفهم فى القراءة الصامتة للصف السادس .

(صورتان ا ، ب) .

سلسلة الاختبارات التحصيلية للمرحلة الابتدائية :

فى اللغة العربية : اختبار لكل صف من الأول إلى الخامس .

فى الحساب : للصف الأول والثانى والثالث (عمليات ومسائل) والرابع (عمليات ومسائل) والخامس (عمليات ومسائل وهندسة) .

فى المواد الاجتماعية : اختبار لكل صف من الثالث إلى الخامس .

فى مبادئ العلوم والتربية الصحية : اختبار لكل صف من الثالث إلى الخامس .

اختبارات الحساب للمدارس الابتدائية :

الصف الأول .

الصف الثانى .

الصف الثالث (عمليات ومسائل) .

بحث التعليم الابتدائى

من أهم البحوث التى قامت بها وزارة التربية والتعليم بالإقليم الجنوبى والتى اشترك فيها المؤلف ، بحث التعليم الابتدائى . ومرجع هذه الأهمية

لا إلى ضخامة الجهود المتضمنة فيه وحسب بل وأيضاً إلى أن التعليم الابتدائى هو القاعدة الكبيرة لكل مراحل التعليم التالية . وهو ، إلى جوار هذا ، يكلف الدولة نفقات طويلة خاصة وأنه بسبيل أن يكون إلزاميا فضلا عن كونه مجانيا . كما أن السنوات التى يقضيها الفرد فى هـذا التعليم من أخطر سنوات تشكيل تكوينه النفسى والانفعالى والجسمى والعقلى . ومن ثم يجب ، بفرض أنه ممكن ، استغلال هذه المرحلة فى إعداد المواطن الصالح .

وقد بدأ البحث فى ١٩٥٥/٣/٢٧ وانتهى التقرير فى سنة ١٩٥٧ وهذه خطواته كما وردت فى تقرير لجنة التعليم الابتدائى .

أولا

الخطوة الأولى : البحث الاستطلاعى.

الخطوة الثانية : البحث التجريبى .

١ ـ استمارة يملؤها المدرس (الصورة التجريبية) .

٢ ـ استمارة يملؤها ناظر المدرسة (الصورة التجريبية) .

٣ ـ استمارة يملؤها مفتش القسم (الصورة التجريبية) .

دراسة تقارير اللجنة الفرعية :

أولا : بيانات إحصائية عن البحث التجريبى فى المناطق الثلاث (القاهرة الشمالية ـ طنطا ـ المنيا) التلاميذ والفصول ـ مستويات المدارس ـ مجموع تلاميذ الفرق المختلفة ـ المبانى ـ نفقات التلميذ .

ثانياً : تفريغ استمارات المدرسين والنظار والمفتشين .

تقرير عن جدول تفريغ استمارات النظار .

ثالثاً : دراسة الاستمارات الثلاث .

١ – تقرير عن استمارة المدرس .

٢ – تقرير عن استمارة الناظر .

٣ – تقرير عن أستمارة مفتش القسم .

استفتاء المدرس (الصورة النهائية) .

استفتاء الناظر (الصورة النهائية) .

استفتاء المفتش (الصورة النهائية) .

كراسة مفتش القسم .

رابعاً : استفتاء الأهالى .

(ا) بطريقة فردية .

(ب) بطريقة الندوات .

البحث فى آراء الأهالى والمواطنين فى مسائل التعليم الابتدائى .

أولا : وجهة نظر الأهالى فى شخصية تلميذ المدرسة الابتدائية :

(عاداته وهواياته – الإقبال على الاطلاع والقراءة – معرفة التلميذ لحقوقه وواجباته – الإلمام بالأحداث العامة الجارية . شخصية الطفل وثقته بنفسه) :

ثانياً : وجهة نظر الأهالى فى المدرسة والتعليم الابتدائى :

(مدى استفادة التلاميذ منه من حيث عدد سنوات الدراسة – اليوم المدرسى والجمع بين التعليم فى نصف اليوم والعمل فى النصف الآخر – استفادة التلاميذ من نظام التغذية – التعليم المختلط « اشتراك البنت مع الولد ») .

ثالثاً : وجهة نظر الأهالى فى أثر المدرسة فى البيئة :

(هل نجحت المدرسة فى إزالة الفوارق بين الطبقات – هل يشترك

المدرسون مع الأهالى فى النهوض بالبيئة – هل امتد أثر المدرسة إلى تثقيف الآباء والأمهات) .

بيانات عامة عن نتائج الاستفتاء الفروى :

أولا : أثر التعليم الابتدائى على التلميذ :

ثانياً : نوع المدرسة المفضل .

ثالثاً : أثر التعليم فى البيئة .

رابعاً : المبانى .

خامسا : المدارس الخاصة بمصروفات .

سادسا : توفير السكن للمعلمين والمعلمات .

سابعا : الأشراف الصحى .

ثانيـــاً

حالة التعليم الابتدائى :

حالة التعليم الابتدائى فى صورة بيانية :

المدارس الابتدائية .

مدارس إعداد معلمى المرحلة الأولى :

المبانى :

الفصول والتلاميذ .

المدارس المستكملة وغير المستكمله .

المدرسون .

الميزانية .

حالة التعليم الابتدائى كما تصورها الاستفتاءات .

أولا : استفتاء مفتش القسم .

ثانيا : استفتاء ناظر المدرسة .

ثالثا : استفتاء المدرس .

حالة التعليم الابتدائى كما تصورها الاختبارات .

أولا : امتحان التحصيل .

ثانيا : امتحان القبول بالمدارس الإعدادية .

التوصيات

أولا : أوجه الإصلاح ووسائله .

ثانيا : علاقة التعليم الابتدائى بالمراحل الأخرى .

ثالثا : الاقتراحات بشأن عمل تخطيط لتعميم التعليم الابتدائى وتحسينه .

رابعا : خاتمة بشأن الاقتراحات الخاصة بالتشريعات والقرارات اللازمة لتنفيذ الإصلاحات التى تقترحها اللجنة .

نماذج الاختبارات المستخدمة فى هذا البحث

الشكل (٢٦)

اختبار العربى للسنة الثانية الابتدائية

الصورة الأولى الفقرة (٤)



يحمل الصياد حمامة

صاد الرجل سمكة

الرجل يأكل السمكة

وهنا يؤشر المفحوص على الجملة التى يصورها هذا الرسم .

الفقرة المكافئة من الصورة الثانية الفقرة (٤)



(شكل ٢٧)

وطبيعى أن يطلب نفس الشيء من المفحوص ؟

الفقرة المكافئة من الصورة الثالثة الفقرة (٤)



(شكل ٢٨)

الفقرة الرابعة من الصورة المكافئة الرابعة

(شكل ٢٩)



اختبار المستوى التحصيلى للسنة الخامسة الإبتدائية :

الصحة :

١ – ضع علامة ✓ أمام العادة الصحية ، وعلامة × أمام العادة غير الصحية فيما يأتى :

تلميذ يكثر من شرب الماء بعد الأكل .

تلميذ يستعمل فرشاة الأسنان بعد الأكل .

تلميذ يجد يديه نظيفة قبل الأكل فيجلس إلى المائدة ويأكل دون غسلهما .

تلميذ يسهر بالليل كثيراً وينام بالنهار .

تلميذ يكثر من أكل المكرونة ويقلل من أكل الخضروات .

تلميذ يجد أظافره طويلة فيقضمها بأسنانه .

تلميذ يفتح شبابيك حجرته وبابها للتهوية ثم يجلس وسطها .
تلميذ يقرأ فى ضوء خافت .

التاريخ :

١ – اختر من المستطيل التالى اسم البناء الذى أقامه كل ملك من هؤلاء الملوك ؟

الملك خوفو شيّد	مدينة الاسكندرية
	منارة الاسكندرية
الملك مينا بنى	مدينة الجدار الأبيض
الملك بطليموس الثانى أقام	الهرم الأكبر بالجيزة
	معبد الدير البحرى

٢ – ضع خطاً تحت الكلمة الصحيحة بين كل قوسين فى الجمل الآتية :

(ا) موقع مدينة منف الآن هو (القاهرة – البدرشين – الأقصر) .

(ب) صنع المصريون القدماء الورق من (قش الأرز – نشارة الخشب – نبات البردى) .

(جـ) اعتقد المصريون القدماء أن الإله الذى سيحاسبهم على أعمالهم يوم البعث هو (رع – أمون – أوزيريس) ؟

(د) يرجع الفضل إلى الملك أحمس فى طرد (بدو الصحراء – الهكسوس – الليبيين) من مصر .

(هـ) انتشرت المسيحية فى مصر فى عهد (الفراعنة – اليونان – الرومان) .

الجغرافيا :

١ ــ أكمل الجمل الآتية :

(ا) من أهم منتجات السودان .. .. و ....

(ب) لون الهولنديين .... أما الهنود فلونهم ....

(ج) أكبر قارات العالم هى ....

(د) الجو فى بلاد الاسكيمو ....

٢ــ (ا) معك بوصلة ، نظرت إليها فوجدت الشمال أمامك ، فأين يقع الشرق بالنسبة لك ؟

....

(ب) كيف يعرف المصريون ليلة أول رمضان من كل عام ؟

....

....

....

....

....

وقد وزع تلاميذ كل فرقة حسب درجاتهم التى حصلوا عليها فى امتحان كل مادة إلى أربع فئات .

(ا) تلاميذ درجاتهم أقل من ٢٥٪ من النهاية الكبرى للدرجات .

(ب) تلاميذ درجاتهم أقل من ٢٥٪ إلى ٤٩٪ من النهاية الكبرى للدرجات

(حـ) تلاميذ درجاتهم أقل من ٥٠٪ إلى ٧٤٪ من النهاية الكبرى للدرجات

(د) تلاميذ درجاتهم ٧٥٪ فما فوق

فكانت نتائج هذه الامتحانات بصفة إجمالية كالآتى :

نتائج الاختبارات التحصيلية

الفرقة	المـادة	جملة عدد الممتحنين	عدد التلاميذ فى كل من الفئات				النسبة المئوية لعدد التلاميذ فى كل فئة			
			ا	ب	حـ	د	ا	ب	حـ	د
الثانية	اللغة العربية	١٧١٤	٢٥٣	٣٨٦	٥١٧	٥٥٨	١٤ر٧٦	٢٢ر٥٢	٣٠ر١٦	٣٢ر٥٦
	حساب	٩١٢	١٠٢	١٩٥	٢٩٧	٣١٨	١١ر١٨	٢١ر٣٨	٣٢ر٥٧	٣٤ر٨٧
الرابعة	اللغة العربية	١٦١٠	٢٤٤	٣٠١	٤٧٢	٥٩٣	١٥ر١٥	١٨ر٧٠	٢٩ر٣٢	٣٦ر٨٣
	حساب	١٠٥٤	٣٠٤	٢٤٨	٣١٠	١٩٢	٢٨ر٨٤	٢٣ر٥٣	٢٩ر٤١	١٨ر٢٢
	اللغة العربية	٥٨٥	٢٣	١١١	٢٥٩	١٩٢	٣ر٣٩	١٨ر٩٧	٤٤ر٢٧	٢٢ر٨٢
السادسة	حساب	٨١٧	٢٤٢	٢٢٢	٢٢٦	١٢٧	٢٩ر٦٢	٢٧ر١٧	٢٧ر٦٦	١٥ر٥٤
	معلومات	٧٢٥	١٢٩	١٦٣	٢٧٨	١٥٥	١٧ر٧٩	٢٢ر٤٨	٣٨ر٣٤	٢١ر٣٨

(جدول ١٥)

كذلك عمل الجدول التالى لتوضيح النسب المئوية لعدد التلاميذ الذين امتحنوا فى كل فرقة ، وفى كل مستوى من المستويات الأربعة .

النسبة المئوية لعدد التلاميذ موزعين حسب مستوى

درجاتهم فى الامتحان التحصيلى لكل مادة

الفرقة	المـادة	النسبة المئوية لعدد التلاميذ الحاصلين على درجات			
		أقل من ٢٥٪	٢٥٪ إلى أقل من ٥٠٪	٥٠٪ أقل من ٧٥٪	٧٥٪ فأكثر
الثانية	اللغة العربية	١٥	٢٢	٣٠	٣٣
	حساب	١١	٢١	٣٣	٣٥
الرابعة	اللغة العربية	١٥	١٩	٢٩	٣٧
	الحساب	٢٩	٢٤	٢٩	١٨
السادسة	اللغة العربية	٤	١٩	٤٤	٣٣
	الحساب	٣٠	٢٧	٢٨	١٥
	معلومات	١٨	٢٣	٣٨	٢١

(جدول ١٦)

ومن هذا نتبين أن من الممتحنين فى مادة اللغة العربية من تلاميذ السنة الثانية .

١٤,٧٦٪ لم يصلوا إلى مستوى الفرقة السابقة (الفرقة الأولى) .

٢٢,٥٢٪ لم يصلوا لأكثر من مستوى الفرقة السابقة (الفرقة الأولى) .

وكذا ٣٠,١٦٪ نجحوا فى تحصيل مستوى الفرقتين (الأولى والثانية) .

٣٢,٥٦٪ منهم ٣٠,١٦ ٪ بدرجة جيدة ، ٢,٤٠ ٪ بدرجة ممتازة .

ولم نرد أن نستطرد فى استعراض نماذج من الاختبارات أو النتائج . إذ أنها ، برغم انتهاء البحث ، ما زالت تحت التجربة فى ضوء التطورات الأخيرة فى نظام التعليم . ويمكن الرجوع إلى مطبوعات الوزارة لمعرفة المزيد .

بحوث المؤلف

قام واشترك وأشرف المؤلف على عدد من البحوث التربوية والاختبارات التحصيلية وأهمها .

١ – بحث تقييم التحصيل فى المدرسة الابتدائية .

٢ – بحث تقييم اختبارات القبول بالمدرسة الإعدادية .

٣ – بحث تتبع خريجى المدارس الصناعية الثانوية .

٤ – بحث تقييم بطاقات الحساب التدريبية التشخيصية .

٥ – إعداد اختبار استانفورد الجديد فى الحساب للاستخدام فى المدارس المصرية بالاشتراك مع الأستاذ إسماعيل القبانى .

٦ – إعداد بطاقات الحساب التدريبية التشخيصية بالاشتراك مع الأستاذ إسماعيل القبانى .

وسنبدأ بعرض بحث التعليم الابتدائى لنفس الأسباب التى سبق ذكرها :

بحث تقييم التحصيل بالمرحلة الابتدائية

فى سنة ١٩٥٥ أشرف المؤلف على بحث لتقييم التحصيل بالمرحلة الابتدائية كان الغرض الأساس منه مقارنة التحصيل فى المناطق والمدارس المختلفة ، إذ يرى البعض أن المناطق القريبة من المدن الكبرى تحظى بإمكانيات وكفايات وإشراف وتوجيه لا يتوفر فى المناطق الأخرى النائية بينما يرى البعض الآخر أن تحصيل التلاميذ خارج المدن الكبرى لا يقل عن تحصيل نظرائهم فيها . ومن هذا كان لا بد من القيام بدراسة أو دراسات موضوعية للإجابة على هذه المشكلة .

خطة البحث

أولا : إعداد الاختبارات : كان من الضرورى إعداد اختبارات موضوعية موحدة تطبق على عينات من تلاميذ المرحلة الابتدائية . تمثل جميع المناطق والمستويات . وتم إعداد أربعة اختبارات موضوعية وضعت لكل منها صورة تجريبية أولى طبقت على عينة تمهيدية . ثم عدلت ووضعت فى صورة تجريبية ثانية وطبقت مرة ثانية على عينة تمهيدية أخرى من تلاميذ المدارس الابتدائية . وعلى ضوء التجربة الثانية ، وما ظهر من تحليل النتائج إحصائيا ، عدلت الاختبارات ووضعت فى صـــورتها النهائية . وفيما يلى وصفها :

١ ــ اختبار فى اللغة العربية : يتكون الاختبار من ٧ أسئلة « فقرات » ينقسم بعضها إلى ثلاث أو أربع أجزاء كما يحتوى الاختبار قطعة للإملاء والزمن المحدد للاختبار كله ٦٠ دقيقة خصص منه ٢٠ دقيقة للإملاء .

٢ ــ اختبار فى الحساب : يتكون من ٢٠ فقرة والزمن المحدد له ٤٥ دقيقة .

٣ ــ اختبار فى المعلومات العامة : يتكون من ٥٠ فقرة ، من نوع فقرات الاختيار من عدة إجابات ، والزمن المحدد له ٣٠ دقيقة .

٤ ــ اختبار للذكاء العام : يتكون من ٤٠ فقرة ، من نوع فقرات الاختيار من عدة إجابات والزمن المحدد له ٢٥ دقيقة .

ثانياً : جمع المعلومات عن التلميذ : أعدت لهذا الغرض بطاقة روعى فيها ذكر المعلومات التى تلقى ضوءاً على نتائج التلميذ فى الاختبارات الموضوعية السالفة الذكر . وقد قام المدرسون بملء البيانات فى هـــذه البطاقات عن كل تلميذ بناءاً على معرفتهم للتلميذ وتقديرهم الذاتى ، كل مدرس فى الناحية التى تخصه . كما خصصت فى البطاقة مسافات لتسجيل

نتائج التلميذ فى إختبارات البحث كذلك خصصت مسافات لتسجيل نتائج امتحان القبول بالمدرسة الإعدادية لمن يتقدم لها من هؤلاء التلاميذ .

ثالثاً : اختيار العينة : تم اختيار ١٤٢٩ تلميذاً (٩٨٥ ولدا ، ٤٣٤ بنتا ، ١٠ لم ترد عن جنسهم بيانات) يمثلون ٢٠ مدرسة تمثل ٩ مناطق تعليمية على الأسس التالية :

(ا) تمثيل المناطق التعليمية المختلفة بالإقليم المصرى كى يمكن تمثيل تلاميذ السنة الرابعة بالمدارس الابتدائية (كانت المرحلة الابتدائية وقتذاك خمسة سنوات) .

(ب) أن تمثل هذه المدارس جميع المستويات الاجتماعية والاقتصادية بحيث تمثل العينة الطبقات العليا والوسطى والدنيا .

(جـ) أن تمثل العينة المدارس الخاصة بالبنين والمدارس الخاصة بالبنات والمدارس المشتركة .

رابعاً : فحص سجلات التلاميذ : لتبين تاريخهم التعليمى : وقد ركزنا على معرفة عدد مرات رسوب كل منهم ، وفى أى عمر رسب ، وما الفرق بين نسب الرسوب فى كل عمر إلى مجموع الراسبين ، وهل هناك فرق فى نسب الرسوب بين الذكور والإناث ، وهل هناك فرق فى نسب الرسوب بين المناطق التعليمية المختلفة . كما كنا نجمع عدد مرات الرسوب على الفرد فى الفرق الدراسية حتى الرابعة الابتدائية سواء فى ذلك من رسب مرة أو مرات ونقارنها بمن لم يرسب إطلاقا . وأيضاً قارنّا بين من لم يرسب ومن رسب مرة واحدة فى أى فرقة وبين من رسب أكثر من مرة طول تاريخه التعليمى :

خامساً : إجراء الاختبارات : أعدت كراسة تعليمات مفصّلة عن طرق إجراء الاختبارات ، وتصحيحها ، وطريقة ملء البيانات فى بطاقة التلميذ .

وأرسلت كراسات التعليمات والاختبارات والبطاقات إلى مفتشى المناطق ليتولوا توزيعها بأنفسهم على نظار المدارس التى اختيرت ضمن عينة هذا البحث . وكان لكل اختبار كراسة تعليمات منفصلة ، كما كان كل اختبار على حدة موضوعا فى كتيب منفصل .

سادساً : وصف العينة : تكونت العينة ، كما سبق القول ، من ١٤٢٩ تلميذاً يمثلون ٢٠ مدرسة تمثل ٩ مناطق تعليمية . ولم يتمكن الباحث من تحديد جنس ١٠ أفراد فى العينة وكان الباقون : ٩٨٥ تلميذاً ، ٤٣٤ تلميذة وكانت أعمار ٦٩٪ من البنين تتراوح بين ٩ ، ١١ سنة بينما كانت أعمار ٧٩٪ من البنات تتراوح بين ٨ ، ٩ سنوات . وكانت النسبة المئوية لمن تزيد أعمارهم عن ١٣ سنة فى العينة الكلية ٥٪ .

سابعاً : النتائج : نقتصر هنا على تقديم وصف إجمالى لنتائج البحث فيما يتعلق بالمستوى التحصيلى . وقد جمعت عن كل تلميذ معلومات عن تقديرات المدرسين له فى اللغة العربية ، والحساب ، والذكاء العام . ولكى نتأكد من فهم المدرسين لما نقصده بالذكاء العام عقدت ندوات اشترك فيها ممثلون لهم كما اشترك فيها مفتشو الأقسام وناقش الباحث معهم هذا المفهوم وكيف يوضع تقرير للتلميذ عنه . كما طبع كتيب منفصل فيه تحديد للمفهوم وبيان لطريقة تقديره تقديراً أقرب إلى الدقة وأرسلت الكتيبات لهم . وبرغم هذا لم نأخذ بتقديرات المدرسين للذكاء إذ أنها لم تتبع تعليماتنا كما لم تكن موضوعة فى مستويات متميزة .

كما حصلنا على درجات أفراد العينة فى اختبارات البحث ، فى اللغة العربية ، والحساب ، والمعلومات العامة ، والذكاء العام . وقد تقدم بعض أفراد العينة لامتحان القبول بالإعدادية فى كل من اللغة العربية والحساب وجمعنا معلومات عن درجاتهم فى هاتين المادتين . وهكذا ، فى تحليل

النتائج ، قسمنا العينـــة إلى مجموعتين إحداهما تقدمت لامتحان القبول والأخرى لم تتقدم فى حين جمعنا بطاقات المعلومات عن تلاميذ كل من القسمين وكذلك درجاتهم فى اختبارات البحث .

ونظراً للاختلاف بين النهايات العظمى لتقديرات المدرسين واختبارات البحث ودرجات امتحان القبول . ولكى يتسنى لنا مقارنة هذه الدرجات حولت جميعها بحيث أصبحت النهاية العظمى لكل منها ١٠٠ . ولما كانت تقديرات المدرسين فى اللغة العربية توضع من درجة عظمى هى ٤٠ فقد ضربت فى ٢½ . ولما كانت درجات اختبارات البحث توضع من نهاية عظمى هى ٨٠ : فإنها كانت تضرب فى ١٫٢٥ : وكذلك ضربت درجات الطلبة فى امتحان القبول للمدارس الإعدادية فى اللغة العربية فى ٢ إذ أن درجتها العظمى كانت ٥٠ . وبنفس الطريقة استعملت معاملات لتحويل باقى الدرجات بحيث تصبح نهاياتها العظمى ١٠٠ وبذلك أمكن مقارنة كل الدرجات معاً :

ثامناً : تبويب الدرجات : بوبت درجات العينة حسب تصنيفها إلى :

١ ــ الفئة التى حسبت لها درجات عن اختبارات البحث كما حصلنا على تقديرات المدرسين لأفرادها .

٢ ــ الفئة التى حسبت لها درجات عن اختبارات البحث ورصدت درجات أفرادها فى امتحان القبول للإعدادية ، إذ أن كل أفرادها تقدموا له ؟

٣ ــ الفئة التى لم يرسب أفرادها أبداً فى سنى الدراسة الابتدائية ؛ وكان بعضهم ينتمى إلى الفئة الأولى والبعض الآخر ينتمى إلى الفئة الثانية :

٤ ــ الفئة التى رسب أفرادها مرة واحدة أو أكثر خلال سنى الدراسة الابتدائية . وقد انتمى بعض أفرادها إلى الفئة الأولى ، وانتمى البعض الآخر إلى الفئة الثانية .

تاسعاً : نتائج البحث :

١ – ارتفعت تقديرات المدرسين ، فى أغلب المدارس ، لأغلب التلاميذ ، كثيراً عن درجات هؤلاء التلاميذ فى كل من اختبارات البحث ودرجات امتحان القبول للإعدادية ؛ وقد تراوحت التقديرات بين ٦٠ ، ٧٠ فى أغلب الحالات ، بينما تراوحت متوسطات الدرجات فى اختبارات البحث وامتحان القبول للإعدادية بين ٤٠ ، ٦٠ . ومن هذا يتضح لنا أن المدرسين ، فى جميع المناطق التعليمية ، بالغوا فى تقدير تحصيل تلاميذهم .

٢ – كان مستوى التلاميذ عموماً فى اللغة العربية ، سواء اعتبرنا اختبارات البحث أو نظرنا إلى درجات امتحان القبول فى هذه المادة ، أعلى من درجاتهم فى الحساب .

٣ – كان هناك اتفاق كبير بين درجات التلاميذ فى اختبارات البحث وامتحان القبول للمرحلة الإعدادية فى كل اللغة العربية والحساب .

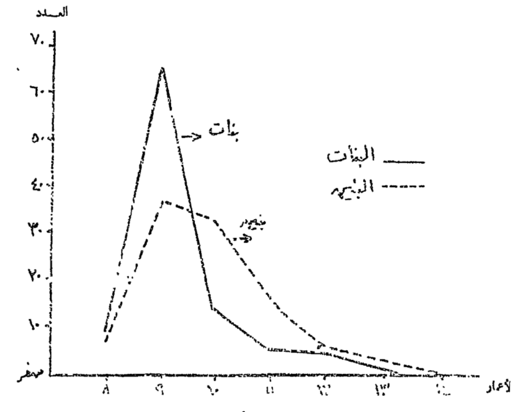
٤ – كانت المتوسطات فى اختبار المعلومات العامة أعلى من المتوسطات فى اختبار الذكاء العام . ولكن الفروق غير ذات دلالة .

٥ – فاق تحصيل البنات فى اختبار المعلومات العامة تحصيل البنين فيه بدرجة كبيرة .

٦ – ظهر أن متوسطات الدرجات تنخفض بتقدم أعمار التلاميذ باستمرار . وكان اطراد الفروق واضحاً ، وإن لم تكن هذه الفروق نفسها كبيرة وخاصة فى اختبارى المعلومات العامة والذكاء العام .

٧ – متوسط أعمار البنات كان أقل كثيراً من متوسط أعمار البنين .

٨ – نسبة الناجحين فى امتحانات النقل انخفضت بتقدم أعمارهم . وقد اتضح هذا ، على وجه أخص ، فيما يتعلق بمن زادت أعمارهم عن ١٢ سنة إذ قلت نسبة نجاحهم فى المتوسط عن ٥٠٪ .



(الشكل ٣٠) توزيع أعمار البنين والبنات
من الشكل يتضح أن أعدادا أكبر من البنات فى الأعمار الصغيرة بينما نجد
أن أعدادا أكبر من البنين فى الأعمار الأكبر

عاشراً : التفسير :

لن يكون تفسيرنا هنا شاملا ، إذ لا يتسع المجال لهذا ، وسنقتصر على ما يمكن أن يعد فروضاً توضع تحت البحث . وسنناقش النتائج مفسرين إياها حتى نخلص إلى الإجابة عن المشكلة الأولى التى دفعت إلى القيام بهذا البحث وهى معرفة ما إذا كان مستوى تحصيل وإعداد تلاميذ المناطق البعيدة عن المدن الكبرى أقل أم أكثر من تحصيل تلاميذ المدارس فى المناطق القريبة من المدن الكبرى أو المدن الكبرى نفسها . ونرى :

١ ــ أن تقديرات المدرسين تزيد كثيراً عن درجات اختبارات البحث ودرجات امتحان القبول لأن المدرسين فى هذا المستوى يفتقرون إلى معرفة

طرق التقدير الدقيقة الموضوعية . كما أنهم ، شعورياً أو لا شعورياً ، يرون أن تقدير أنفسهم وكفايتهم فى التدريس يعتمد على تقدير تلاميذهم الذى إذا ارتفع رفع معه تقدير الآخرين للمدرسين وتقديرهم لأنفسهم .

٢ – ارتفعت درجات التلاميذ فى الحساب عنها فى اللغة العربية لأسباب نقترح أنها قد ترجع إلى عوامل معينة ، ولكن هـذا الاقتراح لا يقوم إلا بدراسة لتحقيقه نرجو أن يحفز هذا البحث البعض لتعهدها . ومن بين هذه العوامل أن اللغة المألوفة المستخدمة فى الحياة اليومية هى اللغة العامية وليست العربية كما تدرس . ومن هنا كان احتمال تمكن التلميذ من اللغة المستخدمة فى الحياة اليومية وميدان العمل فيما بعد برغم ضعفه فى اللغة كما تقيسها الامتحانات أو تصاغ بها الاختبارات ، احتمالا جائزاً . فإذا أضفنا إلى هذا أن مدرسى المدرسة الابتدائية ، وخاصة وقت القيام بالبحث ، غالباً من غير المؤهلين لتدريس العربية تأهيلهم لتدريس الحساب الذى لا يكلفهم خروجاً عن لغة الحياة اليومية الأرقام . ولا ندعى أن هؤلاء المدرسين معدّون لمادة أو لأخرى ، وهم يدرسون كل المواد ، ولكننا نرجح أنهم درسوا أثناء إعدادهم مواد أقرب فى طبيعتها إلى الحساب بل وقد يعتمد بعضها عليه : وما من شك أن تطوير تدريس اللغة فى هذه المستويات الابتدائية يرفع من درجاتها ، ولكن هذه المسألة بدورها تخضع للبحث . ولا ينكر أحد أن التلميذ يتعلم جملا وتراكيباً لغوية لا ترتبط بواقع الاستخدام اليومى ارتباط تدريس الحساب بها ، ونتوقع أن استخدام اللغة العامية مع صور مأخوذة من واقع بيئتنا سيغير الأمور إلى أفضل ، والجزم بقول كهذا غير مشروع دون أن تأيده بحوث نأمل القيام بها .

٣ – اتفاق درجات الاختبارات مع درجات امتحان القبول للمرحلة الإعدادية محك لصدق هذه الاختبارات . وقد يرى البعض أنها ما اتفقت

مع الامتحانات إلا لأنها صيغت صياغتها ، ولكننا ندفع بأنها ما دامت اختبارات تحصيل واختبارات تطبق على التلاميذ فقط وأن البحث تربوى وجب أن تتفق موادها مع محتويات المنهج الدراسى .

٤ — قد يرجع ارتفاع درجات البنات عن درجات البنين إلى أثر البيئة فمن المعلوم أن الأسر التى تذهب بناتها إلى المدرسة أسر من مستوى ثقافى أ... غالباً من الأسر التى تمنعهن عن هذا .

٥ — من الطبيعى أن تنخفض درجات التلاميذ بتقدم أعمارهم وقد يرجع هذا إلى أحد الأسباب التالية :

(ا) أن كبار السن من التلاميذ غالباً ما التحقوا بالمدرسة فى سن متأخر لعدم اهتمام أسرتهم بالتعليم ومن ثم لا يجدون فى ظروفهم ما يدفع إلا المذاكرة ويدعمها .

(ب) كبار السن من التلاميذ فى السنة الرابعة الابتدائية بينهم نسبة كبيرة ممن رسبوا خلال سنى تعليمهم وهذا الأمر يرجح انخفاض ذكائهم أو قلة قدرتهم على التحصيل : هذا فضلا عن أن الرسوب ذاته ، خاصة إذا تكرر ، يثبط هممهم فى التحصيل .

٦ — البنت ، وخاصة بعيدا عن العاصمة ، إما أن تجد تشجيعاً للذهاب إلى المدرسة فى سن مبكرة ، وإما أن تجد رفضاً لمبدأ الذهاب إلى المدرسة . ولهذا فهى إذا ذهبت إلى المدرسة كانت غالباً أصغر سناً من الولد . فإذا أضفنا إلى هذا أن البنات ، وخاصة بعيدا عن العاصمة ، عندما تكبر أعمارهم عن الثانية عشر لا يجدن تشجيعاً من الأسرة للاستمرار فى المدرسة ، علمنا أن من تبقى فى المدرسة غالباً ما تكون صغيرة لم تصل بعد إلى السن الذى تعتبره فيها الأسرة أنثى .

٧ ـ ظهر أن المدى الذى تراوحت فيه الانحرافات المعيارية لدرجات التلاميذ فى كل مادة على حدة أكبر كثيرا من المدى الذى تراوحت فيه الفروق بين متوسطات هذه الدرجات . وبلغة الإحصاء ، كان الانحراف المعيارى للانحرافات المعيارية لدرجات كل مادة أكبر من الانحراف المعيارى لمتوسطات درجات التلاميذ فى هذه المادة .

وقد حدث هذا فى كل المدارس فى كل المناطق القريبة منها من المدن الكبرى وكذا البعيد مما يجعلنا ننتهى إلى أن هناك تفاوتاً كبيرا فى مستويات التحصيل برغم أنها قد تكون ، فى المتوسط ، جيدة أو متوسطة . كما أن حدوث هذا فى كل المدارس فى كل المناطق الأمر الذى يجعلنا نصل إلى أنه لا فرق جوهرى بين تحصيل وإعداد التلميذ فى المناطق القريبة من المدن الكبرى وبينه فى البعيدة عنها وهذا هو جواب مشكلة البحث .

بحث تقييم إمتحانات القبول بالمدارس الإعدادية

قام المؤلف ببحث لتقييم إمتحانات القبول بالمدارس الإعدادية . وقد بدأ البحث سنة ١٩٥٤ وانتهى بتقرير عن نتائجه سنة ١٩٥٦ . وضمنه دليلا باختبارات القبول لهذه المرحلة وما يجب أن يتوفر فيها من خصائص تجعلها جيدة . وإليك مقتطف عن البحث ونتائجه .

دراسة تحليلية لامتحانات القبول بالمدارس الإعدادية

فى مناطق القطر لسنة ١٩٥٤ ـ ١٩٥٥

مقدمة :

تقوم كل منطقة من المناطق التعليمية فى فصل الربيع من كل عام بعقد امتحانين تحريريين فى مادتى اللغة العربية والحساب وذلك للتلاميذ المتقدمين للالتحاق بالمدارس الإعدادية الذين أتموا السنوات الأربع بالمدرسة الابتدائية (هذا ما كان متبعاً وقت القيام بالبحث) .

ويقبل بالسنة الأولى بالمدرسة الإعدادية التلاميذ الحاصلين على ٦٥ درجة أو أكثر فى مجموع المادتين بشرط ألا تقل درجة كل مادة على حدة عن ٢٥ وينقل التلاميذ الحاصلون على مجموع يتراوح بين ٥٠ ، ٦٥ درجة فى المادتين ، بشرط ألا تقل درجة كل مادة على حدة عن ٢٥ ، وينقل هؤلاء إلى فرقة أعلى بالمدرسة الابتدائية (السنة الخامسة) بعد أن يؤدوا امتحانات باقى المواد الأخرى بالمدرسة . أما التلاميذ الحاصلون على أقل من ٥٠ درجة فى مجموع المادتين أو على أقل من ٢٥ درجة فى أية واحدة منهما فيعقد لهم امتحان فى مدارسهم مع بقية التلاميذ الذين لم يتقدموا لهذا الامتحان .

ويتولى المفتشون فى كل منطقة تعليمية وضع امتحانات القبول بالمدرسة الإعدادية والإشراف عليها . ونظراً الى تباين تقدير المفتشين للمستوى الدراسى ، من ناحية ، وإلى تباين القيم الفنية للامتحانات نفسها ، من ناحية أخرى ، يتضح لنا أن توزيع التلاميذ على الفرق لا يخلو من بعض العيوب . فمن الجائز أن بعض التلاميذ قد نقلوا خطأ إلى فرق أعلى بينما حرم من ذلك تلاميذ آخرون ممن هم جديرون بالانتقال . وقد تكون مثل هذه الأخطاء ليست ذات بال بالنسبة إلى موظفى المدرسة ولكنها تعد بالنسبة إلى التلميذ من الخطورة بمكان . أضف إلى هذا أنه إذا كان ينبغى أن يستند توزيع التلاميذ . على نتيجة الامتحانات فإن القياس العلمى يتطلب أن تراعى الدقة فى هذا التوزيع بقدر الإمكان .

لذلك اهتم الباحث بهذه الامتحانات وطريقة إعدادها . فأعد دليلا باختبارات القبول بالمرحلة الإعدادية وضمنه توضيحاً للأغراض التى ينبغى أن يحققها هذا الامتحان والنواحى التى ينبغى مراعاتها عند وضع الامتحان وطريقة التصحيح وتوزيع الدرجات الخ . . . ويتضمن هذا الكتيب أيضاً امتحانين كاملين فى اللغة العربية وامتحانين كاملين فى الحساب ، كأمثلة يهتدى بها عند وضع امتحانات المناطق . كذلك شمل الكتيب طريقة توزيع الدرجات

عن كل سؤال وعن أجزاء السؤال الواحد . وقد وزع هذا الكتيب فى الوقت المناسب على القائمين بإعداد امتحانات المناطق وعقد لهم اجتماع عام بالوزارة نوقشت فيه جميع النواحى التى ينبغى مراعاتها عند وضع امتحانات القبول للمدارس الإعدادية .

واستكمالا لما سبق قام الباحث بالدراسة الحالية التى تهدف إلى الأغراض التالية :

(أ) تحديد مدى الاختلافات الكيفية ، إن وجدت ، بين امتحانات القبول بالمدرسة الإعدادية بوجه عام من منطقة إلى أخرى .

(ب) تحديد مدى الاختلافات الكيفية بين مفردات الامتحان الواحد .

(ج) تحديد الفروق الحقيقية ، إن وجدت ، بين مجموع الدرجات الكمية للتلاميذ من منطقة إلى أخرى ، فى كل اختبار على حدة وفى كلا الاختبارين معاً .

منهج عام

رغبة فى تقويم الامتحانات من الناحية الكيفية ، طلب إلى كل منطقة أن ترسل نسخة من أسئلة الامتحان فى كل من مادتى اللغة العربية والحساب لعام ١٩٥٥ واختير عدد من المحكمين لتقويم هذه الامتحانات .

وقد وصلت إلينا نماذج هــذه الامتحانات من المناطق الخمسة عشر التالية :

أسوان	الفيوم	طنطا وكفر الشيخ
قنا	الجيزة	دمنهور
سوهاج	القاهرة	الإسكندرية
أسيوط	شبين الكوم	المنصورة
بنى سويف	الزقازيق	القناطر

واختارت وزارة التربيـــة والتعليم بعض الأساتذة المختصين لتقويم امتحانات مادة الحساب ، كما اختير لتقويم امتحانات اللغة العربية بعض الأساتذة المتخصصين . وقد طلب إلى كل واحد من المحكمين أن يرتب مضمون الاختبارات فى كل منطقة طبقاً للمعايير الأربعة الآتية :ـ

١ ـ هل السؤال صالح من حيث أنه يقيس عناصر أساسية فى مادة الامتحان ؟

٢ ـ هل السؤال واضح ومحدد ؟

٣ ـ هل السؤال محدود بحيث أن التلميذ يستطيع الإجابة عنه إجابة وافية وقت مناسب ؟

٤ ـ هل السؤال مناسب للمستوى المفروض فى تلاميذ السنة الرابعة ؟

وقد أعطى كل واحد من المحكمين بطاقة يمكنه أن يسجل بها تقديراً عددياً لكل بند من بنود امتحانات المنطقة .

ويبدأ هذا التقدير من ١ ـ ٩ وفقاً لمفتاح التقويم التالى :ـ

(١ و ٢ و ٣) ضعيف

(٤ و ٥ و ٦) متوسط

(٧ و ٨ و ٩) جيد

وإليك نموذج من استمارة التقويم هذه :

استمارة تقويم امتحانات المناطق

اسم الفاحص .... المنطقــة .... مادة الامتحان ....

	رقم الســؤال										
مفتاح التقويم : ضعيف ١ و ٢ و ٣ – متوسط ٤ و ٥ و ٦ – جيد ٧ و ٨ و ٩	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	المجمــوع
١ – هل السؤال صالح من حيث أنه يقيس عناصر أساسية فى مادة الامتحان .											
٢ – هل السؤال واضح ومحدد											
٣ – هل السؤال محدود بحيث أن التلميذ يستطيع الإجابة عنه إجابة وافية فى وقت مناسب .											
٤ – هل السؤال مناسب للمستوى المفروض فى تلاميذ السنة الرابعة											

(ا) ما هو تقديرك العام للاختبار فى مجموعه من حيث ملاءمته لمستوى السنة الرابعة الابتدائية (التقدير كما فى مفتاح التقويم) التقدير العام للاختبار ....

(ب) اكتب فى الفراغ التالى ما تراه من ملاحظات على الامتحان فى ضوء النقط السابقة وغيرها مما تعتقد أن له أهمية بالنسبة للغرض من الامتحان .

(جدول ١٧)

وبعد انتهاء المحكمين من تقدير مفردات الاختبار ، قام كل منهم بتقدير الاختبار ككل تقديراً عاماً مستخدماً فى ذلك مفتاح التقويم السابق .

وقد عمل التحليل الكمى لدرجات التلاميذ فى هذه الاختبارات على عينة تبلغ ١٠ فى المائة من التلاميذ الذين أدوا الاختبار فى كل منطقة .

وعمل توزيع تكرارى للمجموع الكلى للدرجات فى الاختبار ، ولدرجات الاختبار فى مادة الحساب ودرجات الاختبار فى اللغة العربية لكل منطقة .

تقويم الاختبارات تقويماً موضوعياً

أخذ متوسط التقديرات التى وضعها المحكمون (وهم خمسة محكمين لاختبار الحساب ، وسبعة لاختبار اللغة العربية) وذلك للحصول على تقدير مركب لكل بند من بنود الاختبار فى كل منطقة . ثم حسب متوسط هذه التقديرات المركبة للحصول على تقدير جزئى لكل اختبار « بندا بندا » ثم حسب متوسط الانحراف للتقديرات الجزئية هذه للحصول على تقدير مدى تغير المفردات فى كل اختبار . وأخذ أيضاً متوسط التقديرات التى وضعها المحكمون للاختبار ككل للحصول على تقدير عام لكل اختبار . ثم حسب الانحراف فى هذه التقديرات لمعرفة مدى تغير تقديرات المحكمين فى كل اختبار .

ويبين لنا الجدولان (١٨و١٩) التقديرات الجزئية « بندا بندا » والتقديرات الشاملة للامتحان ، والانحرافات المتوسطة ، والمتوسط والانحراف المعيارى فى كل المناطق لاختبارى اللغة العربية والحساب ، وكان معامل الارتباط بين التقديرات الجزئية (بندا بندا) والتقديرات الشاملة هو ٩٥ر٠ لاختبار الحساب و٨١ر٠ لاختبار اللغة العربية . وهذان المعاملان كبيران بحيث يمكن اعتبار أن هذه التقديرات ربما بلغت نفس الدقة التى تسمح بها عمليات التقدير المبنية على أسس موضوعية .

....

ويبين لنا الرسم البيانى رقم (٣١) المتوسط والانحراف المتوسط ومدى التقديرات الذاتية لكل منطقة فى هذين الاختبارين ويتضح من هذا الرسم أن هناك تباينا واسعا بين تقييم المفردات فى بعض المناطق بينما تتقارب كثيرا فى مناطق أخرى .

ولعمل مقارنة عددية بين اختبارات المناطق حسب الدرجات التائيـة (T-scores) للتقديرات الجزئية (بنداً بنداً) وللتقديرات الشاملة لكل منطقة . ويوجد متوسط هذه التقديرات فى العامود الأخير الأيمن بالجدولين رقم (١٨ و ١٩) وفى الرسم البيانى رقم (٣١) .

ويبين الرسم البيانى رقم (٣١) أنه ينبغى مراجعة وتحسين الاختبارات فى مادة الحساب للمناطق الأخيرة الثلاث ولمادة اللغة العربية فى المنطقتين الأخيرتين على الأقل .كما هو الرسم أن المحكمين يرون أن اختبار اللغة العربية كان جيداً بنوع خاص فى اثنتين من المناطق :

ولما كانت امتحانات القبول بالمدارس الإعدادية تلعب هذا الدور الهام فى برنامج التربية ولما كانت هذه الدراسة قد أوضحت تبايناً كبيراً فى هذه الاختبارات ، لذا يتضح وجوب العمل على تهيئة برنامج تدريبى فى وضع الامتحانات لكل من توكل إليهم هذه المهمة من المدرسين وغيرهم حتى يكون هناك تناسق وتكافؤ بين امتحانات المناطق المختلفة ٥

وإليك بعض نتائج البحث موضحة فى جدولين ورسم بيانى :

جدول رقم (١٨)

ملخص التقديرات الموضوعية لاختبارات الحساب

المنطقة	التقدير الجزئى (بنداً فبند) لكل سؤال فى الاختبار		التقدير الشامل للاختبار ككل		الدرجات التائية
	المتوسط	الانحراف المتوسط	المتوسط	الانحراف المتوسط	(T. Scores)
أسوان	٨ر١٦	٠ر٤٠	٨ر٢٠	٠ر٦٤	٥٥
قنا	٨ر٦٤	٠ر٣٠	٨ر٤٠	٠ر٤٨	٦١
سوهاج	٨ر٥٥	٠ر٣١	٨	٠ر٤٠	٥٧
أسيوط	٨ر٣٠	٠ر٣٥	٨ر٤٠	٠ر٤٦	٥٨
بنى سويف	٦ر٣٤	١ر١٥	٦ر٢٠	٠ر٣٢	٢٧ر٥
الفيوم	٧ر٣٢	١ر٢٦	٧	٠ر٨٠	٤٠ر٥
الجيزة	٨ر٠٣	٠ر٨٤	٧ر٤٠	٠ر٤٨	٤٩
القاهرة	٨ر٦٥	٠ر٤١	٨ر٤٠	٠ر٤٨	٦١
شبين الكوم	٧ر٩١	٠ر٥٥	٧ر٤٠	٠ر٤٨	٤٨
الزقازيق	٦ر٧٤	١ر٧٣	٦ر٢٠	٠ر٣٢	٣٠ر٥
طنطا – كفر الشيخ	٧ر٧٦	١ر٢٠	٧ر٦٠	٠ر٤٨	٤٨ر٥
دمنهور	٧ر٩٥	١ر٠٤	٧ر٥٠	٠ر٥٠	٤٩
الإسكندرية	٨ر١٩	٠ر٤٦	٨	٠ر٤٠	٥٤ر٥
المنصورة	٧ر٩٤	٠ر٤٧	٧ر٨٠	٠ر٣٢	٥١
القناطر	٨ر٥٧	٠ر٣٩	٨ر٣٠	٠ر٣٢	٥٩
جميع المناطق	٧ر٩٤	٠ر٦٦ (S. D.)	٧ر٦٥	٠ر٧٠ (S. D.)	

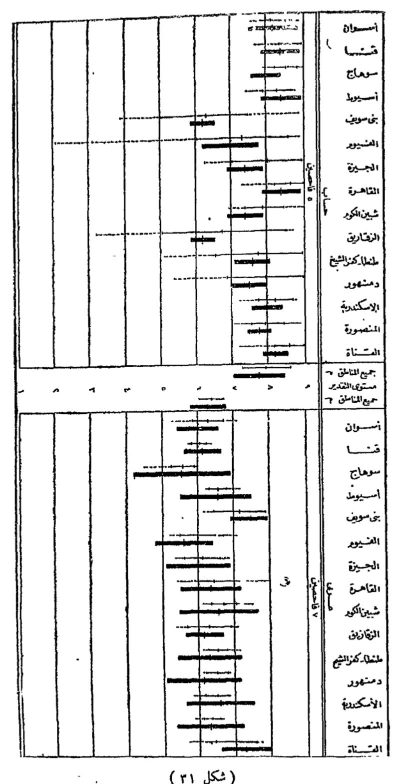
R. = ٠ر٩٥٤ –

الجدول رقم (١٩)

ملخص التقديرات الموضوعية لاختبارات اللغة العربية

المنطقة	التقدير الجزئى بنـدا فبند		التقدير الشامـل (للاختبار ككل)		الدرجات التائية
	المتوسط	الانحراف المتوسط	المتوسط	الانحراف المتوسط	(T.scores)
أسوان	٦ر٣٠	٠ر٤٣	٦ر٠٠	٠ر٥٧	٤٥ر٢
قنا	٦ر٠٦	٠ر١٩	٦ر١٤	٠ر٤٩	٤٣ر٨
سوهاج	٦ر٦١	٠ر٣٣	٥ر٥٧	١ر٣٥	٣١ر٦
أسيوط	٦ر٥٧	٠ر٢٢	٦ر٥٧	ر٩٤	٤٥ر٥
بنى سويف	٧ر١٤	٠ر٧٢	٧ر٤٣	ر٤٩	٧١
الفيوم	٦	٠ر٥٢	٥ر٥٧	٠ر٧٨	٣٧
الجيزة	٦ر١٢	٠ر٥١	٦	٠ر٨٦	٤٢ر٨
القاهرة	٦ر٤٣	٠ر٥٦	٦ر٣٤	٠ر٨١	٥٢ر٧
شبين الكوم	٦ر٧٠	٠ر٣٣	٦ر٥٧	١ر٠٦	٥٦ر٣
الزقازيق	٦ر٢٧	٠ر٦٠	٦ر١٤	٠ر٤٩	٤٦ر٢
طنطا – كفر الشيخ	٦ر٥٦	٠ر٢٧	٦ر٢٩	٠ر٨٨	٤٨ر٤
دمنهور	٦ر٤٨	٠ر٣٣	٦ر١٤	١ر٠٢	٤٨
الاسكندرية	٦ر٥٠	٠ر٣٤	٦ر٥٧	٠ر٩٤	٥٣ر٥
المنصورة	٦ر٨٠	٠ر١٧	٦ر٢٩	٠ر٩٠	٤٩
القناطر	٧	٠ر٥٧	٧ر٢٩	٠ر٧٦	٦٧ر٥
جميع المناطق	٦ر٤١	٠ر٣٧ (S. D.)	٦ر٣٣	٠ر٥٠ (S. D.)	

R = ٠ر٨٠٦



(شكل ٣١)

تقدير المحكمين لامتحانات القبول بالمدرسة الإعدادية ويوضح انخفاض مستوى العربى عموما عن الحساب ؛ وإن كانت الأحكام متقاربة عن العربى أكثر منها عن الحساب إذ نجد أنهم يضعون امتحان الزقازيق فى الحساب فى الرتبة السادسة بينما يضعون امتحان القاهرة مثلا فى التاسعة . والخط الثقيل يمثل الانحراف المعيارى للتقدير العام للاختبار ككل ، والمنقط مداه ، والنقط المنصفة هى المتوسطات . والخط الرفيع يمثل تقدير الاختبار من بنوده

دليل الاختبارات

للقبول بالمرحلة الإعدادية

أعد المؤلف هذا الدليل سنة ١٩٥٥ وأرفق به نماذج لامتحان القبول المقترحة والخصائص التى يجب أن تتوفر فيها . وقد حذت وزارة التربية والتعليم حذوه وأخرجت أدلة مماثلة متتابعة كل عام . ونحن نقتطف لك منه ما يلى :

مقدمة عامة

تنص المادة رقم ٣ من القانون رقم ٢١١ لسنة ١٩٥٣ بشأن تنظيم التعليم الثانوى على أنه يشترط فيمن يقبل بالسنة الأولى من المرحلة الإعدادية أن يؤدى بنجاح امتحان القبول الذى يعقد لهذا الغرض فى مادتى اللغة العربية والحساب .

ومن الواضح أن الغرض من هذا الامتحان هو انتقاء التلاميذ ذوى الاستعداد المناسب لمتابعة الدراسة فى مراحل التعليم الإعدادى .

ولعل أصلح الطرق للانتقاء – لو أمكن الاطمئنان لتوافر الدقة والموضوعية فيها – هى طريقة الاعتماد على الأحكام التى تستخلص من تتبع أحوال التلاميذ طول مدة الدراسة وتسجيل ملاحظات المدرسين عنها . . . ولكن الاطمئنان إلى هذه الأحكام لايتوافر قبل مضى زمن تتحسن فيه أحوال المدارس وتتهيأ للمدرسين أسباب الدراسة الشخصية لكل تلميذ ويكتمل تدريبهم عليه . ولذلك جعل الاختيار فى الوقت الحاضر على أساس امتحان للقبول يدخله كل من أنس فى نفسه حسن الاستعداد له من تلاميذ المدارس الابتدائية مع وضع حدود معقولة للسن حتى يصبح الامتحان سابقة عادلة تتساوى فيها الفرص للحصول على الأماكن الموجودة .

ومهما قيل عن عيوب الامتحان فليس لدينا فى الوقت الحاضر وسيلة عامة أصلح منها للاختيار وهى لاتزال الوسيلة العملية المعترف بها فى العالم أجمع لاختيار الطلاب فى مثل هذه الظروف على أن يراعى فى إعداد أسئلة الامتحانات وطريقة إجرائها وتصحيحها ما يضمن حسن الاستفادة من نتائجها .

ولاشك أن من الخير أن يوجه التلميذ إلى نوع المدرسة المناسبة لاستعداده ومن الظلم أن نقحم طفلا إلى نوع من التعليم لا يستطيع أن يتابعه فى يسر ونجاح ، إننا بذلك نجنى على هذا الطفل ونضيع وقته وجهده ونبتعد به عن التوجيه السليم ، كما نعرضه لمحنة نفسية من الشعور بالضآلة والعجز ، وفى الوقت نفسه نضيع على أولياء الأمور وعلى الدولة مالا وجهدا ونشحن المدارس بمن لا يصلح وبمن يسد الطريق فى وجه من هو خير منه .

لذلك يجب أن يعد هذا الامتحان على نحو يستطاع به تمييز من يستطيع متابعة الدراسة ممن تقف بهم قدراتهم عن المتابعة .

فيجب أن يعد الامتحان بحيث يقيس القدرة على التفكير إلى جانب قياسه القدرة على التحصيل الدراسى ، أعنى أن يقيس الاستعداد كما يقيس التحصيل .

ولكن يجب أن يتم كل ذلك فى بعد عن التكلف والتعقيد حتى يكون الامتحان مقياسا عادلا يقيس فى غير تحيّف أو إعنات من يستطيع أن يواصل الدراسة ممن يحسن — لصالحهم ولصالح الوطن — أن يتجهوا وجهة أخرى.

أولا — اللغة العربية

الغرض من الامتحان :

نسعى فى هذا الامتحان إلى قياس :

١ — قدرة التلميذ على الفهم .

٢ — قدرته على التعبير .

٣ – قدرته على رسم الكلمات رسماً مطابقا لقواعد الهجاء :

٤ – قدرته على الاستعمال اللغوى الصحيح :

٥ – قدرته على الكتابة بخط واضح حسن :

والمقصود بقياس قدرته على الفهم – قياس سرعة القراءة ، وسرعة الفهم وصحته ، مع ما يدخل تحت ذلك من قياس القدرة على الاستنباط ، والربط ، وإصدار الأحكام على المقروء بما يناسب التلاميـــذ فى مثل هذه السن :

كيف نعد مقياساً لاختبار هذه القدرات ؟

أولا – قياس القدرة على الفهم :

الطريقة الأولى :

عرض موضوع مناسب ومناقشته بأسئلة تكشف عن مدى فهم التلميذ لما يقرأ ، وتبدو مناسبة الموضوع للتلاميذ فى النواحى الآتية : –

(ا) الأفكار التى يشتمل عليها الموضوع

(ب) الأسلوب الذى تعرض فيه هذه الأفكار .

(ﺟ) طول الموضوع .

أما من حيث الأفكار :

فيراعى ألا تكون أفكاراً معنوية بحتة ، كما يجب أن يدور الموضوع حول خبرات وتجارب يستطيع التلميذ أن يتمثلها ويتصورها ويحكم عليها ؛ ويجب أن تكون هذه الأفكار واضحة مرتبة ترتيباً مفهوما .

وأما من حيث الأسلوب :

فيجب تحاشى الغريب من الألفاظ والعبارات ، كما يجب تحاشى الإسراف فى تطويل الجمل ، واستخدام أساليب التقديم والتأخير ، والجمل الاعتراضية ، وكل ما من شأنه أن يعجز التلميذ عن متابعة القراءة فى يسر وفهم .

وأما من حيث الطول :

فيراعى أن تكون القطعة فى الحدود التى تعود التلميذ أن يراها فى كتب المطالعة التى بين يديه ، وتتراوح عدد كلمات القطعة المناسبة لهذا الامتحان بين مئة كلمة ومئة وعشرين .

صياغة الأسئلة :

ومن حيث صياغة الأسئلة التى يناقش بها الموضوع ، وتختبر بها القدرة على الفهم فيحسن أن يكون لكل سؤال غرض ، وأن يجس كل سؤال ناحية تكشف عن جانب من الفهم وتناقش القطعة ببضعة أسئلة :

١ - بعضها مباشر ، إجابته مذكورة بنصها فى الموضوع .

٢ - وبعضها غير مباشر ، تستنبط إجابته من القطعة .

٣ - وبعضها أسئلة الغرض منها قياس قدرة التلميذ على إصدار حكم على المقروء ، كالحكم على شخص أو شىء بصفة معينة ، أو قياس قدرته على فهم الموضوع فهما كليا .

الطريقة الثانية :

وتكون على نحو مما يأتى :

(ا) بعرض قصة قصيرة مناسبة تنقصها بضع كلمات ، ويطلب من التلميذ أن يقرأ القصة وأن يضع فى كل مكان خال الكلمة المناسبة له من

بين الكلمات المكتوبة له فى أعلى القصة .

(ب) أو بعرض قصة قصيرة فى عبارات غير مرتبة ويطلب من التلميذ أن يرتب هذه العبارات ترتيبا يؤلف منها قصة متصلة .

وينبغى أن تكون هذه العبارات مناسبة فى طولها وعددها .

الطريقة الثالثة :

ويقصد بها اختبار مدى قدرة التلميذ على فهم مدلول اللفظ ، وما يقابله أو يماثله أو قدرته على اختيار لفظ بعينه لا تتم الجملة بدونه ، كأن يطلب منه ذكر عكس معنى كلمة مذكورة له أو اختيار كلمة من بين ما يماثلها ، أو تكميل جملة بلفظ من نوع معين كأن يكون ضميرا ، أو اسما موصولا ، أو اسم إشارة .

ثانيا — قياس القدرة على التعبير :

والأنواع السابقة من اختبار قدرة التلميذ على الفهم تقيس كذلك قدرته على التعبير إلى جانب قياس قدرته على الفهم ؛ لكنه تعبير قصير ومقيد بالإجابة عما هو مطلوب منه .

أما قدرته على التعبير المنطلق فيمكن أن تقاس بمطالبته بالكتابة فى موضوع أو رسالة .

(أ) الكتابة فى موضوع :

يطلب إلى التلميذ أن يتحدث فى موضوع مناسب ، فى حدود قدرته وخبرته ومشاهداته ويراعى أن يكون هذا الموضوع منتزعا من البيئة التى يعيش فيها :

١ ـ فابن القرية الذى يعيش بين الحقول ويرى الفلاح فى عمله ويشاهد الماشية والدواب ويرى عن كثب الطيور والدواجن غير ابن المدينة :

٢ ـ وابن المدينة غير ابن العاصمة التى تعج من حوله بالحيـــاة ، ووسائل المواصلات والراديو ودور السينما والحدائق العامة والآثار والمعارض والحفلات .

٣ ـ وهناك الحياة المدرسية التى قد تزدحم بمظاهر النشاط من رحلات وحفلات وألعاب وتربية طيور فى حظائر . . .

من هذه الأودية وما يشبهها ومن غيرها مما يتصل بحياة التلميذ ويتناسب مع بيئته وظروفه يمكن أن تشتق الموضوعات التى يطلب إلى التلميذ الكتابة فيها .

كيف يُصاغ عنوان الموضوع :

ويحسن أن تترك للتميذ حرية الانطلاق فى التعبير ، فلا يقيد بعناصر أو تفصيلات يتضمنها رأس الموضوع مما قد يقيد انطلاقه ، إلا إذا كان ذلك ضروريا لمساعدته على الكتابة .

وكذلك يحسن أن يترك التلميذ حرا يكتب ما شاء ، قل ذلك أو كثر ، فلا يقيد بسطور محدودة كما جرى على ذلك بعض الممتحنين ، إذ ليست العبرة بكثرة ما يكتبه التلميذ أو بقلته إنما العبرة بالكيف لا بالكم .

على أن تقييد التلميذ بعدد محدود من الأسطر ليس تيسيرا له إنما هو تقييد له وتضييق عليه .

(ب) الكتابة فى رسالة :

ويمكن أن يطلب إلى التلميذ الكتابة فى موضوع مما أشير إليه سابقا ، ولكن فى صورة رسالة كأن يطلب إلى تلميذ من أبناء القرية أن يكتب

خطابا إلى صديق له بالقاهرة يدعوه إلى زيارة القرية وسيذكر له ما سيراه فيها .

ويراعى كذلك أن يترك التلميذ – فى كتابة الرسالة – حرا ينطلق فى تعبيره كما شاء دون أن يقيد بعناصر معينة ، أو بعدد محدود من الأسطر ، إلا إذا كان ذلك ضروريا لمساعدته على كتابة الرسالة كما قدمنا .

تقدير إجابة التلاميذ فى التعبير :

التلميذ الذى يتقدم لهذا الامتحان طفل صغير السن ، محدود الثقافة ، قليل الخبرة والتجربة .

وهذه أمور يجب أن نضعها نصب أعيننا ، ونحن نتخير له موضوع التعبير ، كما ينبغى أن نتنبه إليه ونحن نراجع عمله وإجابته ، ولكن ليس معنى ذلك أن نتقبل منه الغث التافه ، ولكن نكتفى :

١ – بأن يكون تعبيره وضحاً مفهوما .

٢ – وأن تكون أفكاره ، منوعة ، ونقصد بالتنوع هنا أن يشتمل ما يكتبه على فكرتين أو ثلاث ، ولا يقف عند فكرة واحدة غامضة .

٣ – وأن يكون تعبيره سليما فى حدود طاقة مثله ، فلا نتطلب منه أن يكتب بأسلوب عربى فصيح ، بل سنراه يمزج باللغة العربية اللغة الدارجة ، ولكن يجب أن يكون نصيب العربية أغلب وأن يكون طابعها أوفى ، فأدوات النفى والاستفهام والأمر والنهى وأسماء الإشارة والموصولات وما إليها مما درسه فى المرحلة الابتدائية – يجب أن تكون عربية ولكن يتسامح معه حينما يذكر كلمة عامية مكان كلمة عربية كأن يقول « شنطة » بدلا من « حقيبة » مثلا :

ثالثاً – قياس القدرة على الاستعمال اللغوى :

الذى يعنينا هنا هو قياس قدرة التلميذ على استعمال ما درس من أدوات وأساليب ولذلك يجب على واضع الأسئلة أن يراجع بدقة المنهج الذى درسه التلاميذ فى مادة التدريب اللغوى وأن تدور أسئلته حول ما درسوه .

ويجب أن نقيس هـــذه القدرة فى مواقف تشبه تلك المواقف التى تستخدم فيها اللغة فى الحياة . فمن أجل ذلك يحسن أن تدور الأسئلة حول موقف فى قصـــة أو موضوع سهل قصير حتى لا تبدو الأسئلة فى جو صناعى متكلف .

رابعاً – قياس القدرة على الهجاء :

وتختبر هذه القدرة بإملاء موضوع أو قصة ، ويجب أن يكون هذا الموضوع من حيث أفكاره وأسلوبه ومستوى صعوبة كلماته معنى وهجاء بحيث يستطيع التلميذ أن يتابعه فى يسر .

ويجب ألا نتعمد اختيار كلمات غريبة المعنى أو تزدحم بأنواع الهمزات وألوان الصعوبات الإملائية بل ينبغى أن تكون القطعة فى مستوى ما نجده فى كتب القراءة التى بين أيدى التلاميذ وكلماتها مستوحاة من قاموسهم وفى مستوى ما مر عليهم من قواعد الهجاء فى السنوات الماضية .

أما من حيث طول القطعة فينبغى أن تتراوح كلماتها بين ٦٠ و ٧٠ كلمة .

وتقسم القطعة حين الإملاء إلى مقاطع مناســبة ، تملى على التلميذ مقطعا مقطعا .

خامساً – قياس القدرة على الخط :

جرت العادة أن يخصص سؤال للخط . يطالب فيه التلاميذ بكتابة عبارة معينة ، مرة بخط النسخ ، ومرة أخرى بخط الرقعة . ونرى أن التلاميذ فى

هذه السن لا يستطيعون أن يميزوا تمييزاً دقيقاً بين خط النسخ وخط الرقعة حين يكتبون أو أن يقدروا الاختبار فى مادة الخط تقديراً ، ولذلك فهم غالباً يكتبون العبارة التى جاءت لقياس مدى قدرتهم على حسن الخط بمثل الخط الذى يكتبون به الإجابة عن سائر الأسئلة .

ولذلك نرى ألا يخصص سؤال للخط ولكن يخصص له عشر درجات من المئة وتقدر درجة التلميذ على أساس ما يلاحظ من حسن خطه ووضوحه أو رداءته وغموضه فى ورقة الإجابة .

ثانياً ــ الحساب

أولا ــ تتلخص أغراض هذا الاختبار فيما يلى :

١ ــ الهدف التحصيلى : يهدف هذا الإختبار إلى معرفة مستوى تحصيل التلميذ فى مادة الحساب فيما درسه فى المرحلة الأولى ، أى مدى إتقانه فى العمليات الحسابية الأساسية وفروعها المختلفة ومدى فهمه للمسائل التى تحتاج إلى خطوات مختلفة ، يزداد عدد هذه الخطوات تبعاً لازدياد صعوبة المسألة . وتختلف أنواعها تبعاً لدرجـــة ارتباط أجزاء المسألة بهدفها النهائى أى أن للاختبار هدف معيارى يقيس المستوى الذى وصل إليه الطفل .

٢ ــ الهدف التنبوئى : ويهدف الاختبار أيضاً إلى معرفة المستوى العقلى للتلميذ فى الحساب وفى القدرات المختلفة التى ترتبط ارتباطاً وثيقاً بالاستعدادات التى تساعد على التحصيل الرياضى للمراحل التعليمية المقبلة أى الكشف عن المواهب المختلفة فى باكورة تفتحها والتى لها علاقة وثيقة بدراسة الرياضة فى المرحلة الإعدادية والمراحل التى تليها .

٣ ــ الغرض التعليمى : ويجب أن يكون للاختبار قيمة فى تحسين طرق التدريس بطريق مباشر أو غير مباشر .

ثانياً — أهم النواحى التى يقيسها الاختبار ما يلى :

١ — المهارة فى إجراء العمليات الحسابية : ويقصد بها السرعة فى إجراء هذه العمليات فالتلميذ قد يجرى العمليات بسرعة ولكنه يخطئ فيها أو قد يحلها حلا صحيحاً ولكنه يأخذ وقتاً أطول مما يجب . ومهمتنا التأكد من اكتساب التلميذ للمهارة التى تقوم على السرعة والدقة معاً فى إجراء العمليات ، وهذه المهارة من الأحداث الهامة لتدريس الحساب لأنها الأساس الذى يقوم عليه حل المسائل وتقوم عليه دراسة الرياضة فيما بعد . وتحتوى العمليات على الأعداد الصحيحة ، والكسور الاعتيادية والكسور العشرية . ولا شك فى أن تحديد زمن الاختبار يعطى فكرة عن المهارة فى إجراء تلك العمليات .

٢ — معرفة العلاقات الخاصة بالمقاييس والموازين وغيرها : إن التلميذ لا يستطيع أن يحل مسألة على المقاييس أو الموازين أو المكاييل أو غيرها إلا إذا فهم المسألة أولا وعرف العلاقة بين الوحدات المختلفة فى المسألة ومعرفة هذه الحقائق جزء من تحصيله فى الحساب ولذلك وجب علينا أن نقيس مدى معرفته بهذه العلاقات .

٣ — القدرة على الاستنتاج : ليست دراسة الحساب مجرد اكتساب مهارة فى إجراء العمليات الأصلية لكنها تعمل على زيادة تدريب التلاميذ على الاستنتاج وهذا يتصل بالذكاء العام ولذلك فإن قياسها يعطينا فكرة عن مدى صلاحية التلميذ لمتابعة نوع من الدراسة التالية كما يميز بين التلاميذ فى قدرتهم على الاستنتاج .

٤ — القدرة على الفهم : يلعب الدور الرئيسى فى كل خطوة من خطوات التدريس سواء أكان فى دراسة الحقائق أو غيرها أو فى حل المسائل وإجراء

العمليات ولذلك يجب التميز بين التلاميذ فى قدرتهم على الفهم . ومن أمثلة ذلك فهم النظام العددى بإدراك العلاقة بين الأعداد ومعرفة الخانات التى تشغلها الأرقام فى العدد ، وفهم معنى الكسور بأنواعها وفهم معنى الاصطلاحات الحسابية .

٥ ــ القدرة على حل المسائل : الملاحظ فى دراسة الحساب بوجه خاص أن التلاميذ قد يستطيعون إجراء العمليات الأصلية ومع ذلك يتعثرون فى المسائل ويرجع ذلك إلى النظر إلى الحساب على أنه علم آلى يهتم بالمهارة فى إجراء العمليات فقط وتأتى المسائل كالتطبيقات على هذه العمليات وهنا يجب التميز بين التلاميذ فى قدرتهم على الفهم والاستنتاج وحل المسائل واستخدام المهارة فى هذا الحل . وقد سبق التحدث عن الفهم والاستنتاج ، أما القدرة على حل المسائل فإنها تقاس بإعطاء مسائل ليحلها التلاميذ .

ثالثاً ــ طريقة عمل الاختبار :

إن طريقة عمل اختبار هى التطبيق المباشر لما سبق التحدث عنه وهى مرحلة فى غاية من الأهمية وغاية الخطورة ولذلك يجب على واضع الاختبار أن يكون يقظاً وأن يختار العناصر بغاية الدقة لأن المسألة ليست مجرد إعطاء عناصر إنما يجب أن يكون لكل عنصر غرض واضح وأن يكون مستواه من حيث الصعوبة ومن حيث التعقيد مناسباً لمستوى التلاميذ من حيث تحصيلهم الدراسى ونموهم العقلى ، ولهذا يجب أن نسترشد بالنقط التالية فى عمل الاختبار .

١ ــ كثرة الأسئلة وتدرجها : يراعى أن تكون الأسئلة كثيرة وقصيرة ومتدرجة فى الصعوبة والغرض من كثرة الأسئلة أن تتنوع بحيث يشمل الاختبار النواحى السابقة التى يجب أن يقيسها الاختبار فلا يقتصر على حل المسائل فقط وإنما يشمل أشياء ذكرت فيما سبق كالمهارة ومعرفة العلاقات . « الخ »

ولا ننسى أن يكون طول الاختبار مناسباً للزمن المحدد له ولمستوى التلاميذ من حيث التحصيل والنمو العقلى ولهذا فإن المهارة فى إجراء العمليات على الأعداد الصحيحة مثلا تتطلب إعطاء أمثلة على الجمع والطرح والضرب والقسمة . وإذا طبقنا ذلك على الكسور الاعتيادية والكسور العشرية لكان عدد عناصر الاختبار كثيرا جداً ولذلك يحسن أن نكتفى بمثالين فقط فى الأعداد الصحيحة فنأخذ مثلا الجمع والقسمة فى حالة الأعداد الصحيحة لأن الضرب عبارة عن جمع متكرر ، والقسمة عبارة عن طرح متكرر لهذا يحسن أن نأخذ مثلا من الجمع وآخر من القسمة أو من الطرح وثالث من الضرب ؛ وأخذ عنصرين فى كل من الأعداد الصحيحة والكسور الاعتيادية والكسور العشرية يعطينا ست عناصر وهى ملائمة لقياس المهارة . وليس هذا فحسب بل يجب العناية باختيار العناصر من حيث درجة الصعوبة فإن مثلا بسيطاً جداً على الجمع يكون عديم الفائدة أو قد يفقد الاختبار قيمته لأن التلميذ سبق أن درس ذلك فى المراحل المتقدمة جداً والمفروض أن يكون قد بلغ مرحلة أعلى من هذا المستوى البسيط . وكذلك يحسن إختيار نوع الجمع فى جمل فيها ثلاثة أعداد مع عدم تساوى عدد الأرقام فى الأعداد ، هذا يعطى سؤالا صعباً نوعاً ما من أسئلة الجمع .

أما فى القسمة يحسن اختيار نوع ليس سهلا كل السهولة بل فيه صعوبة نسبية . مثل وجود الصفر فى خارج القسمة . والسبب فى ذلك أن تتناسب أسئلة الاختبار مع المستوى الذى نتوقع أن يصل إليه التلاميذ .

أما فى الكسور الاعتيادية فيحسن اختيار سؤال لجمع الأعداد الكسرية مع عدم تساوى المقامات حتى لا يكون السؤال بسيطاً جدا ولا صعباً جدا فيخطىء فيه أغلب التلاميذ . وينطبق نفس العمل على الكسور العشرية فنجعل الخانات العشرية غير متساوية . . . الخ ؛

٢ ــ تنوع الأسئلة : يلاحظ دائماً التنوع فى وضع الأسئلة لقياس النواحى المختلفة فى المهارة فلا تقتصر على قياس عملية واحدة أو عمليتين فقط بل تشتمل على العمليات الأربع بصورها . لذلك يجب ألا يكون عدد الأسئلة قليلا بحيث نجد أن السؤال الواحد يشمل أفكارا كثيرة تغطى المنهج كله فى مسألتين أو ثلاثة ، بل يجب أن تتدرج الأسئلة فى صعوبتها بحيث تشمل أغلب أبواب المنهج ملمة بالأشياء الرئيسية التى نريد قياسها . وفى نفس الوقت يؤدى هذا التدريج فى الصعوبة إلى التميز بين التلاميذ المختلفين . فالتدريج فى الصعوبة مهم حتى لا يفاجأ التلاميذ بمسائل صعبة قد تسبب ارتباكهم وخطأهم مما قد ينعكس على قدرتهم فى حل المسائل التالية فيجب أن نبدأ بالأسئلة التى يحتمل أن يحلوها ليكتسبوا شيئاً من الثقة فى نفوسهم ويعملوا على حل الاختبار .

٣ ــ اتصال المسائل بالبيئة : يجب أن تراعى كل منطقة اختيار المسائل المتصلة بالحياة العملية وبالبيئة المحلية وخاصة فى حفظ علاقات بين وحدات القياس فمثلا تهتم البيئات الريفية بالمكاييل والموازين أكثر من اهتمام البيئات المدنية وهكذا .

٤ ــ النوع المطلوب قياسه فى الأسئلة : يجب أن يشمل الاختبار على أسئلة لقياس النواحى الهامة فى تدريس الرياضة وقد سبق ذكرها . ويهمنا هنا أن نوضح نوع الأسئلة وكيفية اختيارها فمثلا لإختيار المثال الخاص بالقدرة على الاستنتاج ، نختار (٣٦ ÷ ٤ = ٤ + ٠٠٠) وهى أعداد بسيطة . أما اختيار أعداد كبيرة فإنه يجعل المسألة صعبة صعوبة تدخل فيها المهارة فى إجراء العمليات فمثلا يجب عدم (استخدام ٤٦٥ ÷ ١٥ = ٢٠٠ ــ ٠٠٠) أو ما يشابه ذلك ، هذا لأن الحد الأول هنا مشكلة فى ذاته وكذلك الحد الثانى وليس الغرض من هذا السؤال قياس المهارة إنما هو قياس القدرة على الاستنتاج من العلاقات لأن للمهارة أسئلة أخرى سبق التدريب عليها .

ولذلك يحسن أن تكون الأعداد فى حدود الجداول المحفوظة .

وكذلك الأسئلة التى تطلب اختيار الأعداد التى تتمم سلسلة من الأعداد بعد معرفة العلاقة بينها . ومن الأمثلة فى هذا المجال (اشطب العدد الذى لا يتمشى مع سلسلة الأعداد ٣ ، ٦ ، ٩ ، ١١ ، ١٢ ، ٠ ٠ ٠) مثـــلا .

ومن الأسئلة ما يقيس الفهم لمعنى النظام العددى والأعداد الصحيحة لمعرفة القدرة المكانية للرقم فى العدد وفى هـــذا المجال يمكن أن تعطى أمثلة كما يلى :

١ ــ اكتب أكبر عدد مكون من رقمين

٢ ــ اكتب أصغر عدد مكون من رقمين دون ستخدام الصفر

٣ ــ اكتب أكبر عدد مكون من رقمين مختلفين

٤ ــ اكتب أصغر عدد مكون من رقمين مختلفين دون استخدام الصفر.

وهكذا فى حدود معقولة تناسب التلاميذ فلا تطلب منهم مثلا أكبر عدد مكون من خمسة أرقام مختلفة أو أربعة أرقام يكون أصغرها فى خانة كذا . . ؟ أى يجب أن نتجنب محاولة الإيقاع بالتلميذ وإعطائه مشاكل لا داعى لها .

٥ ــ تقدير الدرجات :

من عوامل نجاح الاختبار فى قياس مدى تحصيل التلاميـــذ لمادة الحساب أن توضع درجات مناســـبة لكل نوع من أنواع الأسئلة الموضوعة ، فمثلا :

العمليات الحسابية : يصح أن يخصص لها ٣٠٪ تقريباً من الدرجة النهائية من ورقة الامتحان .

(ا) يحتاج معلم القراءة إلى معرفة المستوى الذى بلغه كل تلميذ من تلاميذه حتى يستطيع أن يساعد الضعفاء ، ويرشد المتوسطين والمتفوقين . وهذا الاختبار يمده بالمعلومات الدقيقة التى تمكنه من توجيه جهوده توجيهاً رشيدا فى هذه النواحى .

(ب) يحتاج معلم القراءة إلى معرفة ما حصله تلاميذه حتى يطمئن إلى سلامة الطريقة التى اتبعها فى تعليمهم أو عدم سلامتها : وهذا الاختبار أداة دقيقة تحقق له هذا الغرض .

(ج) يحتاج معلم القراءة أو المشرف على تعليمها إلى تجريب طريقة جديدة للتعليم ، أو مادة جديدة للقراءة ، ويريد أن يعرف أثر هذه الطريقة أو المادة فى تحصيل التلاميذ . وهذا الاختبار أداة تجريبية دقيقة يستطيع أن يقيس بها هذا الأثر .

(د) يحتاج المشرفون على التعليم ، سواء كانوا نظارا أو مفتشين أو مديرى مناطق ، إلى معرفة التقدم الذى أحرزته مدارسهم أو أقسامهم أو مناطقهم فى هذه الناحية الأساسية من نواحى التعليم فى المرحلة الابتدائية ، وهى ناحية القراءة : وفى هذه الحالة يلزمهم أداة موضعية تمدهم بنتائج يمكن الاعتماد عليها فى تقويم العمل وتوجيهه . وهذا الاختبار يحقق لهم ذلك .

(هـ) يحتاج الباحث فى ميدان التربية عامة ، وفى ميدان تعليم اللغة خاصة ، إلى أداة لقياس القدرة على القراءة ، سواء كان ذلك لتجريب طريقة جديدة فى التعليم ، أو للمقارنة بين عدة طرق ، أو لإدخال تنظيم معين فى المنهج ، أو لتحديد العلاقة بين قدرات دراسية مختلفة ، أو لغير ذلك من الأغراض : وهذا الاختبار أداة دقيقة تعين الباحث على تحقيق هذه الأغراض :

ثانياً : تكوين الاختبار ونواحي القراءة التي يقيسها :

ينقسم هذا الاختبار قسمين : القسم الأول ، ويتكون من ثلاثة أجزاء : (١) اختبار القدرة على تعرف المفردات ، (٢) اختبار القـــدرة على تعرف الجمل ، (٣) اختبار القدرة على فهم معنى الجملة ؛ القسم الثانى ، ويقتصر على اختبار القدرة على فهم معانى الفقرة . والنواحى الأربع التى يتكون منها القسمان نواح أساسية فى القدرة على القراءة ؛ إلا أن الناحيتين الأوليين تبرز أهميتهما فى الصفوف الأولى ثم تأخذ فى التلاشى بعد ذلك ، بينما تبقى الناحيتان الأخيرتان طوال جميع مراحل التعليم .

والاختباران الأولان يشتملان على أسئلة مصورة ، ويبدآن بمواد يسهل على التلاميذ المتوسطين فى الصف الأول أن يميزوها ، وتنتهى بمواد تتطلب درجة كبيرة من القدرة على التمييز . والاختباران الأخيران يشتملان على عدد من الجمل والفقرات ، ويبدآن بجمل وفقرات يسهل على التلميذ المتوسط فى الصف الأول أن يفهمها ، وينتهيان بجمل وفقرات يصعب فهمها على التلميذ المتوسط فى الصف الرابع .

وهذه الاختبارات كلها اختبارات قدرة ، وليست اختبارات سرعة . ومعنى ذلك أنها تقيس ما يستطيع التلميذ أن يعمله فى كل ناحية من هـــذه النواحى إذا أتيح له الوقت الكافى للقراءة .

ثالثاً : حدود الاختبار :

هذا الاختبار يقيس مدى من القدرة يمتد من نهاية الصف الأول إلى نهاية الصف الرابع الابتدائى . وهو مناسب لقياس المتفوقين من تلاميذ الصف الأول ابتداء من منتصف العام الدراسى ، ثم إنه يميز تمييزا دقيقاً بين تلاميذ

الصفين الثانى والثالث ، ولكنه لا يميز بنفس الدرجة بين تلاميذ الصف الرابع وبخاصة فى أواخر العام الدراسى .

(ا) تعليمات عامة :

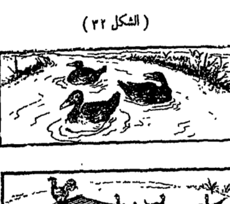
١ — هذا الاختبار مقسم إلى قسمين ، لكل منهما كراسة خاصة ؛ فعند إجرائه يجب تخصيص جلسة لكل قسم حتى يجد التلاميذ الوقت الكافى للإجابة ، ويبذلوا أقصى ما فى وسعهم ؛ ولكن يجب أن تكون الجلستان متقاربتين فى الزمن ؛ ويحسن أن يجرى الاختبار بقسميه فى يوم واحد لضمان عدم تغيب بعض التلاميذ فى القسم الثانى .

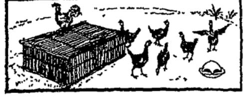
ويمكن الرجوع إلى كراسة تعليمات الاختبار لمعرفة باقيها . وإليك جزء من تعليمات الجزء الثانى من القسم الثانى .

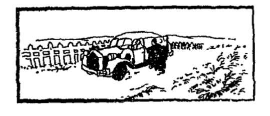
(ب) الجزء الثانى : تعرف الجمل :

١ — ابدأ هذا القسم بأن تقول : « بص فى صفحة ٩ . الصفحة دى فيها تلات صور ؛ وتحتهم جملة مكتوبة . اقرأ الجملة وافهمها » (انتظر قليلا حتى ينتهوا من القراءة) . بص فى التلات صور اللى فوق الكتابة ، ونقى منهم الصورة اللى توافق الكلام المكتوب فى الآخر ، وقول لى أنهى صورة فيهم : الأولى ، ولا التانية ، ولا التالتة ؟ (اترك لهم فرصة للإجابة ، ثم ناقشهم حتى يصلوا إلى أنها الصورة الأولى) . بعد ذلك قل : « وعشان كده حنحط جنبها على اليمين علامة صح . (ارسم على السبورة بروازاً يماثل برواز الصورة ، ثم اكتب على يمينه علامة صح) . بعد ذلك قل : « كل واحد منكم يمسك القلم ، ويحط جنب صورة البط على اليمين علامة صح » ؛ (مر بينهم لتتأكد من أنهم عملوا المطلوب » .

(الشكل ٣٢)







البط فى المساء

فقرة من اختبار سرس اليان فى القراءة الصامتة

تفسير نتائج الاختبار :

يعطى المفحوص درجات عن فقرات الاختبار تبعاً لمفتاح التصحيح ثم تحول هذه الدرجات إلى درجات جيمية تقابلها مستويات . وإليك الجدول رقم (٢٠) .

ثم تحسب الدرجة التائية المقابلة للدرجة الخام باستخدام الجدول رقم ٢١ الذى نأخذ منه القطاع التالى .

(جدول رقم ٢٠) الدرجات الخام وما يقابلها من الدرجات الجيمية والمستويات

الصف الأول			الصف الثانى			الصف الثالث			الصف الرابع		
الدرجة الخام	الدرجة الجيمية	المستوى	الدرجة الخام	الدرجة الجيمية	المستوى	الدرجة الخام	الدرجة الجيمية	المستوى	الدرجة الخام	الدرجة الجيمية	المستوى
صفر	٢	ضعيف جدا	صفر — ٥	٢	ضعيف جدا	صفر — ٩	صفر	عاجز جدا	صفر — ٢١	صفر	عاجز جدا
١ — ٢	٣	ضعيف	٦ — ١٠	٣	ضعيف	١٠ — ١٤	١	عاجز	٢٢ — ٢٩	١	عاجز
٣ — ٥	٤	تحت المتوسط	١١ — ١٥	٤	تحت المتوسط	١٥ — ١٨	٢	ضعيف جدا	٣٠ — ٣٥	٢	ضعيف جدا
٦ — ٨	٥	متوسط	١٦ — ٢٠	٥	متوسط	١٩ — ٢٣	٣	ضعيف	٣٦ — ٣٨	٣	ضعيف
٩ — ١١	٦	فوق المتوسط	٢١ — ٢٥	٦	فوق المتوسط	٢٤ — ٢٨	٤	تحت المتوسط	٣٩ — ٤١	٤	تحت المتوسط
١٢ — ١٥	٧	جيد	٢٦ — ٣٠	٧	جيد	٢٩ — ٣٣	٥	متوسط	٤٢ — ٤٣	٥	متوسط
١٦ — ١٩	٨	جيد جدا	٣١ — ٣٦	٨	جيد جدا	٣٤ — ٣٧	٦	فوق المتوسط	٤٤ — ٤٥	٦	فوق المتوسط
٢٠ — ٢٥	٩	ممتاز	٣٧ — ٤١	٩	ممتاز	٣٨ — ٤٢	٧	جيد	٤٦ — ٤٧	٧	جيد
٢٦ — فأكثر	١٠	ممتاز جدا	٤٢ فأكثر	١٠	ممتاز جدا	٤٣ — ٤٧	٨	جيد جدا	٤٩ — ٥٠	٨	جيد جدا
						٤٨ — ٥٠	٩	ممتاز			

جدول رقم (٢١) للدرجات الخام للاختبار والدرجات التائية
المقابلة لها فى كل صف من الصوف الأربعة

الدرجة الخام	الدرجة التائية				الدرجة الخام	الدرجة التائية			
	الصف الأول	الصف الثانى	الصف الثالث	الصف الرابع		الصف الأول	الصف الثانى	الصف الثالث	الصف الرابع
صفر	٣٦ر٥	٢٣	١٨	—	٢٦	٧٢ر٥	٥٨	٤٥	٣٠
١	٣٨ر٥	٣٤	١٩	—	٢٧	٧٣ر٥	٥٩	٤٦	٣٠ر٥
٢٣	٧٠ر٥	٥٥	٤٢	٢٨	٤٩	—	٨٠	٦٩	٦٤
٢٤	٧١	٥٦	٤٣	٢٨ر٥	٥٠	—	٨١	٧٠	٦٦ر٥
٢٥	٧٢	٥٧	٤٤	٢٩ر٥					

ثبات الاختبار وصدقه :

حسب معامل الثبات بتنصيف الاختبار إلى نصفين وتطبيق معادلة سبيرمان — براون . وكانت معاملات الثبات كما يلى :

جدول رقم (٢٢) معاملات ثبات الاختبار فى الصفوف
الثلاثة الأخيرة من المرحلة الابتدائية

الصف	معاملات الثبات				
	تعرف المفردات	تعرف الجمل	فهم الجملة	فهم الفقرة	المجموع
الثانى	٠ر٧٨	٠ر٥٨	٠ر٧١	٠ر٨٢	٠ر٩٢
الثالث	٠ر٧٧	٠ر٨١	٠ر٧٢	٠ر٨٤	٠ر٩٢
الرابع	٠ر٧٥	٠ر٨٦	٠ر٨٠	٠ر٨٤	٠ر٩٢

أما عن صدق الاختبار فيمكن أن يستدل عليه بارتباطه بمقياس خارجى. والمقياس الذى اختير لمعرفة ارتباطه بالاختبار هو تقدير المدرسين ، فقد أجرى الاختبار على ٥٠٠ تلميذ وتلميذة موزعين على ١٣ مدرسة ابتدائية من مدارس مديرية المنوفية ، وطلب من المعلمين فى هذه المدارس تقدير درجات لهؤلاء التلاميذ والتلميذات فى القدرة على القراءة . ثم حسب معامل الارتباط بين نتائج الاختبار وتقديرات المعلمين فكان ٧٥ر٠ وهو ارتباط عال نسبياً يدل على أن الاختبار يقيس فعلا قدرة اتفق جمهور المعلمين على تسميتها بالقدرة على القراءة .

ويمكن الاستدلال على صدق الاختبار أيضاً بمعرفة العلاقة بين زيادة درجات الاختبار ونمو التلاميذ من صف إلى صف . والجدول رقم (٢١) يشتمل على الدرجات التائية للاختبار فى كل صف من الصفوف الأربعة . وبمقارنة الدرجات التائية فى كل صف من الصفوف بعضها ببعض ، وبمقارنة الدرجات الخام المقابلة للدرجة التائية ٥٠ فى كل صف من الصفوف يتضح أن هذا الاختبار يقيس قدرة تزداد بانتظام من صف إلى صف :

الحساب

من الخطأ أن يعتقد البعض أن تدريس مادة بعينها أو منهج بذاته هدف من أهداف التربية والتعليم . والأمر على عكس ذلك تماماً لأن تدريس أى مادة أو منهج هو وسيلة أو أداة لتحقيق الأهداف العامة للتربية :

ومن أهم الأهداف العامة للتربية تقديم خدمات للفرد فى حياته الخاصة . وأهمها المهارات التى تمكنه من أداء مطالب الحياة وأداء الواجبات التى يفرضها عليه المجتمع الذى يعيش فيه . والفرد الناجح فى حياته هو الذى يستطيع التعامل مع الآخرين فى سهولة ودقة والتعامل فى المجتمع الحديث يقوم على إتقان العمليات الحسابية أى الدقة فى حساباته وكذلك على السرعة فى إجراء

هذه العمليات . لذلك كان اكتساب المهارة فى إجراء العمليات الحسابية لتحقيق السرعة والدقة فى مقدمة أهداف تدريس الحساب .

وكبرى المشكلات التى تعترض تحقيق هذه الأهداف هى أنها تحتاج إلى مران وتدريب طويلين مع العناية بناحيتى الدقة والسرعة معا .

وكثيراً ما ساد بين المشتغلين بالتربية اعتقاد خاطئ مؤداه أن هدف تعليم مادة الحساب هو القدرة على حل المسائل الحسابية المعقدة . ولذلك كثر ابتكار مشاكل ومسائل غاية فى التعقيد . هذا الاعتقاد هو امتداد لما كان يراه المشتغلون بالتربية قديما من أن المران على حل مشاكل ومسائل من هذا النوع يقوى التفكير عند الفرد ويساعده على حل المشاكل المعقدة التى تقابله فى شتى ميادين الحياة .

ولكن الدراسات التربوية التجريبية الحديثة أثبتت بطلان هذا . وقد ارتكز إثباتها على الأسباب التالية :

أولا : أنه لا يحدث انتقال لأثر التدريب من حل المشكلات الحسابية كما تدرس إلى حل ما عداها من المشكلات الأخرى .

ثانياً : أن هذا النوع من المسائل المعقدة لا يصادف الفرد فى حياته العملية بهذه الصورة : على أن الهدف من تدريس الحساب لا يقتصر على مجرد المعرفة بالطرق والقواعد ، ولكن الهدف الأهم أن يكتسب الفرد مهارة فى أدائها فى سهولة ويسر أى تحقيق أكبر مستوى من الدقة والسرعة . وقد ثبت تجريبياً أن المران والتدريب هو الطريق الوحيد إلى تكوين عادات الدقة والسرعة . والمران والتدريب المستمر على عمليات الحساب شاق وممل لكل من التلميذ والمدرس ، فالتلميذ يسأم التكرار ولا يجد الدافع القوى الذى يتغلب على أثر الملل والسآمة ، وكذلك المدرس لا يتسع وقته ولا ترتاح نفسه

لأن يراجع ويصحح كل ما يؤديه تلاميذ فرقة أو فصل دراسى من التدريب .

ولكن هذا التدريب يجب أن يكون فى العمليات التى لا يجيد أداءها التلميذ بدرجة عالية من الدقة أو التى لا تصل سرعته فيها إلى حد مقبول . لذلك يكون من الضرورى تحديد العمليات التى يحتاج فيها التلميذ إلى التدريب سواء لتحقيق الدقة أو لتحقيق السرعة . أى أنه يجب أولاً أن نقوم بتشخيص أنواع العمليات التى لا يتقنها التاميذ والتى لم يصل فى آدائها إلى سرعة مناسبة .

وقد أعد الأستاذ إسماعيل القبانى مع المؤلف مجموعة من البطاقات سميت (بطاقات الحساب التدريبية التشخيصية) تتكون من عدة مجموعات وتتكون كل مجموعة من عدد من البطاقات من الورق المقوى . والمجموعة الواحدة تخص نوعا من العمليات الحسابية . والبطاقة الأولى من كل مجموعة هى اختبار تشخيصى يتناول جميع الاحتمالات الرئيسية المتعلقة بنوع العمليات التى تتناولها هذه المجموعة . وتسهيلا للاستعمال خصص لكل مجموعة من البطاقات لون يميزها عن المجموعات الأخرى . وفيما يلى جدول يبين المجموعات والعمليات التى تتناولها وعدد البطاقات فى كل منها .

وعلى أحد وجهى كل بطاقة طبعت عمليات حسابية وأمام كل منها مسافة يكتب التلميذ فيها الإجابة . وعلى الوجه الآخر طبعت الإجابات الصحيحة بحيث تنطبق تماما على المسافات المخصصة للإجابة فى الوجه الأول . (فى شكل ٣٣) صورة لوجهى البطاقة ز — ٢

وقد أعدت كراسة للتدريب من الورق الشفاف ، بحيث يضع التلميذ البطاقة التى سيعمل فيها تحت الورقة الشفافة الأولى من كراسة التدريب يسجل إجاباته عليها بقلم الحبر . وعند انتهاء الزمن المحدد للتدريب يقلب الوجه الآخر للبطاقة (الإجابة) ويدخله تحت الورقة الشفافة التى كتب

رقم الاختبار **(ز – ٢)**

لجنة التأليف والترجمة والنشر

اختبارات الحساب التدريبية

(طرح – الكسور الاعتيادية)

أجر عمليات الطرح الآتية

$\frac{٢}{٣} - \frac{١}{٢} =$	$\frac{٤}{٥} - \frac{٢}{٥} =$	$\frac{٩}{١٠} - \frac{١}{١٠} =$
$\frac{٧}{٨} - \frac{١}{٢} =$	$\frac{٢}{٣} - \frac{١}{٤} =$	$\frac{٣}{٥} - \frac{١}{٣} =$
$\frac{٧}{١٦} - \frac{١}{١٦} =$	$\frac{٥}{٦} - \frac{٢}{٣} =$	$\frac{١٥}{١٦} - \frac{٣}{٤} =$
$\frac{١}{٢} - \frac{٣}{١٠} =$	$\frac{٩}{١٢} - \frac{١}{١٢} =$	$\frac{٣}{٤} - \frac{٢}{٣} =$
$\frac{٣}{١٠} - \frac{١}{٤} =$	$\frac{٥}{٦} - \frac{١}{٥} =$	$\frac{١٣}{١٦} - \frac{١}{٤} =$
$\frac{٥}{٦} - \frac{٧}{١٢} =$	$\frac{٣}{٤} - \frac{٣}{١٦} =$	$\frac{٥}{٨} - \frac{٥}{١٢} =$
$\frac{٢}{٥} - \frac{٣}{١٠} =$	$\frac{٤}{٥} - \frac{١}{٥} =$	$\frac{٢}{٥} - \frac{١}{٥} =$
$\frac{١}{٤} - \frac{١}{١٠} =$	$\frac{٧}{٨} - \frac{٣}{١٦} =$	$\frac{١١}{١٢} - \frac{٥}{٨} =$

أجوبة الاختبار رقم (ز – ٢)

اسم التلميذ .... الفرقة ....

التاريخ .... رقم الاختبار ....

$\frac{١}{٦}$	$\frac{٢}{٥}$	$\frac{٤}{٥}$
$\frac{٣}{٨}$	$\frac{٥}{١٢}$	$\frac{٤}{١٥}$
$\frac{٣}{٨}$	$\frac{١}{٦}$	$\frac{٣}{١٦}$
$\frac{١}{٥}$	$\frac{٢}{٣}$	$\frac{١}{١٢}$
$\frac{١}{٢٠}$	$\frac{١٩}{٣٠}$	$\frac{٩}{١٦}$
$\frac{١}{٤}$	$\frac{٩}{١٦}$	$\frac{٥}{٢٤}$
$\frac{١}{١٠}$	$\frac{٣}{٥}$	$\frac{١}{٥}$
$\frac{٣}{٢٠}$	$\frac{١١}{١٦}$	$\frac{٧}{٢٤}$

المحاولة رقم ...... عدد العمليات الصحيحة .... الرتبة ....

ممتاز (م) ٢٣ جيد (ج) ٢٠ ناجح (ن) ١٨

(شكل ٣٣) وجه البطاقة (العمليات) ظهر البطاقة (مفتاح التصحيح)

جدول رقم (٢٣)

المجموعة	لون المجموعة	الموضـــوع	عدد الاختبارات	عدد البطاقات
هـ	أزرق	تغيير صورة الكسر	٥	٩
و	رمادى	جمع الكسور الاعتيادية	٦	٦
ز	أصفر	طرح الكسور الاعتيادية	٤	٤
ح	أخضر	ضرب الكسور الاعتيادية	٦	٦
ط	أحمر	الكسور العشرية	٧	٧
ى	رمادى	ضرب الأعداد العشرية وقسمتها	٥	٧
ك	أخضر	تحويل النسب المئوية إلى صور كسرية وبالعكس	٦	٦
ل	أصفر	تحويل النسب المئوية إلى صور كسرية وبالعكس	٧	٧
م	رمادى	المتوسط الحسابي	٥	٥
ن	أزرق	المقاييس المألوفة	٩	٩
س	أصفر	الربح	٧	٧
الفحص العام	أخضر	اختبار عام فيما سبق	٣	٦
		المجموع	٧٠	٧٩

محتويات بطاقة الحساب التدريبية التشخيصية

عليها حلوله فيجد الإجابات الصحيحة ظاهرة بجوار إجاباته التى سجلها . وبذلك يقوم التلميذ بنفسه بتصحيح إجاباته ثم يقوم بجمع عدد الصحيح منها ويسجل هذا العدد فى المكان المخصص لذلك بنفس الصفحة . وفى نهاية كل بطاقة يوجد بيان للمستويات (ممتاز — جيد — ناجح) ليقيم التلميذ نفسه على أساسها . فإذا كان ناجحاً أو جيداً أو ممتازاً ينتقل فى اليوم التالى إلى البطاقة التالية . أما إذا لم يحصل على درجة النجاح كان معنى ذلك أنه لا زال محتاجاً إلى التدريب على هـذا النوع من العمليات . ولذلك يستمر فى التدريب على نفس البطاقة مرة أو مرات ولا ينتقل من بطاقة إلى أخرى إلا إذا نجح فى الأولى .

وفى كراسة التدريب صفحات لتسجيل العمل اليومى وصفحات مبسطة لعمل رسم بيانى « بالأرقام » يساعد التلميذ على معرفة مدى تحسنه أو تقدمه .

وقد أعدت البطاقات جميعها بحيث تعطى زمناً واحداً ثابتاً هو ٤ دقائق للإجابة عن أى بطاقة . وعلى هذا يخصص المدرس ٧ دقائق تقريباً من أول كل حصة حساب يمارس التلاميذ فيها جميعا تدريبهم وتسجيلهم للنتائج . وكل يعمل بمفرده وفقاً لمستواه وسرعته . فقد نجد التلميذ الأول يعمل على الاختبار (و — ٢) من المجموعة (و) عن جمع الكسـور الاعتيادية بينما يعمل التلميذ الثانى على الاختبار (ز — ٤) عن طرح الكسور الاعتيادية وهكذا .

التدريب بهذه الطريقة يبسط عملية التدريب ويختصر وقتها بالنسبة للمدرس والتلميذ على حـد سواء . كما أنه يتميز بالدفع الطيب إذ يتبارى كل فرد أساساً مع نفسه لأنه يقارن تقدمه يوماً بيوم ، وتأتى منافسته مع غيره فى المرتبة الثانية .

وفيما يلى وصف لطريقة استعمال هذه البطاقات كما وردت فى كراسة التدريب :

طريقة استعمال البطاقات

اتبع هذه التعليمات بدقة

١ - عندما تسمع التنبيه بالاستعداد للتدريب الحسابى ، أزل كل ما على درجك من كتب وأدوات ، ثم أخرج كراسة التدريب وقلماً ونشافة .

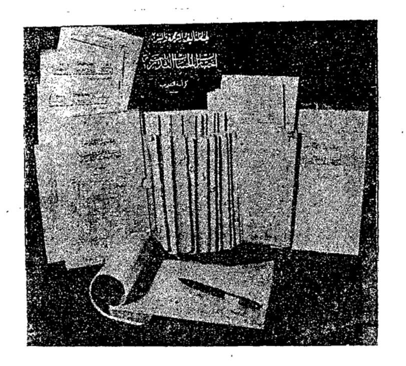
٢ - افتح كراسة التدريب وانظر فى سجل العمل اليومى (من ص ٥ إلى ص ١٠) لتعرف رقم الاختبار الذى ستقوم بإجرائه .

٣ - خذ بطاقة هذا الاختبار ، ولاحظ أن كل بطاقة مكونة من وجهين ، الأول : يحتوى على الأسئلة وتجد دائماً فى ركنه العلوى من اليمين مستطيلا فيه حرف المجموعة ورقم الاختبار ، والوجه الثانى فيه الأجوبة ، ولا تحتاج إليه إلا بعد الانتهاء من الاختبار عندما يأذن لك المدرس بذلك .

٤ - ضع البطاقة تحت أول ورقة شفافة فى الكراسة ، بحيث يكون الوجه الذى به الأسئلة أمامك ، وتكون الحافتان السفلى واليمنى منطبقتين تماما على حافتى الكراسة . ثم اقفل الكراسة وانتظر الأمر ببدء الاختبار .

٥ - فى بعض الأحيان يكون الاختبار شاغلا وجهى البطاقة ، وفى هذه الحالة ستجد فى نهاية الصفحة من أسفل كلمتى « له بقية » ، وعند الوصول إليهما اقلب الوجه الثانى ، ثم ضعه تحت الورقة الشفافة الثانية من كراستك ، واستمر فى الإجابة عن الأسئلة بدون توقف . أما أجوبة هذا الاختبار فتجدها فى بطاقة ثانية ، وعليك عند التصحيح أن تضع صفحة الأجوبة الأولى تحت الورقة الأولى ، والصفحة الثانية (تابع أجوبة . . .) تحت الورقة الثانية .

شكل (٣٤)



مجموعة كاملة من بطاقات الحساب التدريبية التشخيصية
وأمامها كراسة التدريب مفتوحة .

٦ — عندما يقول المدرس « ابتدئ » ، ابدأ فى الإجابة على الاختبار ، واكتب الإجابات بالحبر فى المواضع المخصصة لها بالضبط مع جعل الأرقام واضحة ، اجتهد فى أن تشتغل بسرعة ، ولكن مع مراعاة الدقة التامة حتى لا تقع فى خطأ . استعمل هامش الورقة الشفافة للتسويد إذا لزم :

٧ — عندما يقول المدرس « ضع القلم » ، ارسم خطا رأسيا بعد آخر عملية كنت تجريها ، ثم ضع القلم من يدك مباشرة .

٨ — عندما يقول المدرس « ابتدئ فى التصحيح » ، اقلب البطاقة (أو خذ بطاقة الأجوبة فى الاختبار الذى من صفحتين) ثم ضعها تحت الورقة الشفافة ، وبذلك تصبح الإجابات الصحيحة بجوار إجابتك (على يمينها غالباً) :

٩ — املأ البيانات الموجودة فى رأس هذه الصفحة (الاسم — التاريخ — الفرقة — رقم الاختبار) .

١٠ — ضع علامة × على الإجابات الخاطئة فقط ، واترك الصحيحة بدون علامة . وتعتبر الإجابات خاطئة فى الحالات الآتية :

(ا) إذا كانت الأرقام غير واضحة .

(ب) إذا كان أى رقم قد صحح أو كشط .

(حـ) إذا كان أى رقم مختلفاً عن الجواب المطبوع .

١١ — اكتب رقم المحاولة أمام كلمة (المحاولة رقم . . .) التى فى أسفل الصفحة . فمثلا إذا كانت هذه ثالث مرة تتدرب فيها على اختبار معين فاكتب رقم ٣ .

١٢ — احسب عدد الإجابات الصحيحة واكتبه فى المكان المخصص له :

١٣ — قارن بين هـذا العدد والأعداد المقابلة للرتب الموجودة فى أسفل الصفحة ، ثم اكتب الحرف الدال على رتبتك فى هذه المحاولة أمام كلمة (الرتبة . . .) .

١٤ — دوّن هذه النتيجة بالحبر فى جدول العمل اليومى بهذه الكراسة ، وكذلك اكتب عدد جميع الأسئلة التى أجبت عنها (الخطأ والصواب معاً) فى الخانة الخاصة بذلك فى الجدول المذكور .

١٥ — انزع الورقة الشفافة التى كتبت عليها أجوبة الاختبار وسلمها للمدرس .

١٦ — إذا كنت غائباً فى أحد أيام التمرين فاكتب كلمة غائب أمام التاريخ بالجدول اليومى . وإن كنت حاضراً ولكنك أعفيت من التدريب ، أو قضيت الوقت فى التمرين الحر ، فاكتب « حاضر » . وفى كلتا الحالتين لا تكتب رقماً لليوم .

١٧ — لا تنتقل من اختبار قبل أن تنجح فيه تماماً : لأن الاختبارات متدرجة فى الصعوبة ، ويتوقف كل اختبار على ما قبله . ويكفى للنجاح أن تحصل على رتبة ناجح (ه) المبينة فى كل اختبار ، ولـكن لا تقنع بذلك إذا كان فى مقدورك أن تحصل على رتبة ممتاز أو جيد فإذا لم تحصل على الرتبة المطلوبة فى أول مرة فـكرر التدريب على الاختبار نفسه إلى أن تصل إليها .

١٨ — إذا كانت أخطاؤك كثيرة يمكنك أن تنقل رءوس العمليات التى أخطأت فيها أو التى لم تستطع حلها ، ثم تتمرن عليها أو على عمليات أخرى من نوعها فى ورقة خارجية بالمنزل ، حتى تتقن معرفتها ويتسنى لك حلها بسرعة .

١٩ — إذا أجريت اختباراً ثلاث مرات متتالية دون أن تتقدم فيه بالمرة ، فانتقل إلى الاختبار الذى بعده ، وعد للاختبار الأول بعد مدة .

٢٠ ــ إذا لزم لك فى اختبار أكثر من ثلاث محاولات ، ولاحظت أنك تتقدم ، فابدأ بعمل رسم بيانى لذلك الاختبار على الصفحة المخصصة له (انظر نموذج الرسم البيانى فى صفحة ٦٧٧ (شكل ٣٥) .

٢١ ــ إذا نجحت فى الاختبار التشخيصى الموجود فى أول أى مجموعة ، دل ذلك على أنك لاتحتاج إلى إجراء بقية اختبارات هذه المجموعة ، وبذلك توفر الوقت وتنتقل إلى المجموعة التى بعدها . وللنجاح فى الاختبار التشخيصى يجب ألا يزيد الخطأ عن عملية واحدة فى جميع الاختبار . وإذا لم تنجح فيه لأول مرة فإنه يجوز لك أن تجريه مرة ثانية إذا كان أملك فى النجاح كبيراً .

سجل العمل اليومى

رقم اليوم	التاريخ	رقم الاختبار	رقم المحاولة	عدد عمليات الاختبار	عدد العمليات المحلولة	عدد العمليات الصحيحة	الرتبة

الفحص العام :

والمجموعة الأخيرة من البطاقات التدريبية التشخيصية مخصصة للفحص العام . فقد أعدت بحيث تعطى تشخيصاً دقيقاً لجميع العمليات الحسابية التى تناولتها البطاقات التدريبية جميعها . فبعد أن يجيب التلميذ على اختبارات الفحص العام ، وعددها ثلاث ، يحدد العمليات التى أخطأ فيها ثم يرجع إلى جدول الفحص العام ، المطبوع على غلاف هذه المجموعة ، فيجد فيه أمام كل سؤال (عملية) أرقام البطاقات التى يجب التدريب عليها إذا كان التلميذ أخطأ فى هذا السؤال . وإليك صورة لجدول الفحص العام :

ما يدلنى عليه الفحص العام

إذا أخطأت فى العملية رقم	فإنه يجب علىّ أن أتدرب على الاختبارات	إذا أخطأت فى العملية رقم	فإنه يجب علىّ أن أتدرب على الاختبارات
١	ه : ٢ ، ٣ ، ٤ — و : ١ ، ٢ ، ٣ ، ٤	١٧	ك : ١ ، ٣ ، ٤ ، ٥
٢	ه : ٢ ، ٣ ، ٤ — و : ١ ، ٢ ، ٣ ، ٤ ، ٥	١٨	ك : ١ ، ٢ ، ٣ ، ٤
٣	ه : ٢ ، ٣ ، ٤ — ز : ١ ، ٢	١٩	ك : ٣ — ل : ١ ، ٢ ، ٤ ، ٦
٤	ه : ١ ، ٢ ، ٣ ، ٤ — ز : ١ ، ٢ ، ٣	٢٠	ك : ٢ — ل : ٤ ، ٥ ، ٦
٥	ه : ٢ — ح : ١	٢١	ك : ٢ — ل : ٤ ، ٥ ، ٦
٦	ه : ٢ ح : ٢	٢٢	ك : ٢ — ل : ١ ، ٢ ، ٣ ، ٦
٧	ه : ٢ — ح : ١ ، ٢	٢٣	م : ١ ، ٢ ، ٣ ، ٦
٨	ه : ١ ، ٢ — ح : ١ ، ٢	٢٤	م : ٢ ، ٣ ، ٥
٩	ه : ٢ — ح : ١ ، ٣	٢٥	م : ٢ ، ٣ ، ٤
١٠	ه : ١ ، ٢ - ح : ١ ، ٢ ، ٣ ، ٤	٢٦	م : ٢ ، ٣ ، ٦
١١	ط : ١ ، ٣ ، ٤ ، ٥	٢٧	م : ٢ ، ٣ ، ٤ ، ٥
١٢	ط : ١ ، ٣ ، ٦	٢٨	ك : ١ - م : ٦
١٣	ط : ١ - ى : ١ ، ٢	٢٩	ك : ١ — م : ٦
١٤	ط : ١ — ى : ٣ ، ٤	٣٠	ك : ٢ — م : ٦
١٥	ه : ٣ — ل : ٤	٣١	م : ٧ ، ٨ - م : ٥
١٦	ك : ١ ، ٢ ، ٥		

نمـوذج لملء
سـجل العمل اليومى

رقم اليوم	التاريخ	رقم الاختبار	رقم المحاولة	عدد عمليات الاختبار	عدد العمليات المحلولة	عدد العمليات الصحيحة	الرتبة
١	٨ – ١٠ – ٥٩	٥ – ١	١	٣٠	١٩	١٩	—
٢	٩		٢		٢١	١٩	—
٣	١٠		٣		٢٥	١٨	—
٤	١١		٤		٢٥	٢١	—
٥	١٢		٥		٢٨	٢٦	[illegible]
١٨	٤ – ١١ – ٥٩	٥ – ٤	١	٣٣	٢٣	٢٢	
١٩	٥		٢		٢٤	٢١	
٢٠	٦		٣		٢٩	٢٦	
٢١	٧		٤		٣٠	٢٩	[illegible]

كيفية عمل الرسم البيانى

لعمل الرسم البيانى لأى اختبار اعرف من الفهرس الصفحة المخصصة لذلك ، ثم سر فى العملية بالطريقة الآتية :

بعد كل محاولة ضع نقطة على عدد الأسئلة التى أجبت عنها فى العمود الخاص بالمحاولة . ثم ارسم (بالمسطرة) مستقيما بين هذه النقطة ونقطة عدد الإجابات فى العمود الذى قبله مباشرة . وبنفس الطريقة ضع نقطة على عدد العمليات التى أجيب عنها إجابة صحيحة فى العمود الخاص بالمحاولة ثم ارسم

(بالمسطرة) مستقيما منقطاً بين هذه النقطة والنقطة المقابلة لها فى العمود السابق . وأخيراً ارسم مستقيما أفقياً على العدد الكلى للعمليات فى الاختبار ومده فى كل محاولة .

وإليك مثال لكيفية عمل الرسم البيانى :

رقم الاختبار																
								هـ-٤				هـ-١				
المحاولة رقم	المحاولة رقم	المحاولة رقم	المحاولة رقم	المحاولة رقم	المحاولة رقم	المحاولة رقم	المحاولة رقم	المحاولة رقم ٤	المحاولة رقم ٣	المحاولة رقم ٢	المحاولة رقم ١	المحاولة رقم ٥	المحاولة رقم ٤	المحاولة رقم ٣	المحاولة رقم ٢	المحاولة رقم ١
٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥	٣٥
٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤	٣٤
٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣	٣٣
٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢	٣٢
٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١	٣١
٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠	٣٠
٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩
٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨	٢٨
٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧
٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦
٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥
٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤
٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣	٢٣
٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢
٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١
٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠	٢٠
١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩
١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨	١٨
١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧	١٧
١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦	١٦
١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥	١٥
١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤	١٤
١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣	١٣
١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢	١٢
١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١	١١
١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠

(شكل ٣٥) نموذج لإحدى صفحات الرسم البيانى فى كراسة التدريب

وقد قام المؤلف بتجربة لتقييم بطاقات الحساب التدريبية التشخيصية هذه . وفى التجربة ، طبق اختبار ستانفورد الجديد فى الحساب فى إحدى المدارس الإبتدائية : وبعد معرفة مستويات فصول المدرسة ، فى الحساب ، اخترنا ثلاث فصول من السنة السادسة الإبتدائية متجانسة المستوى . وكان فصلان . منها مجموعة ضابطة والثالث المجموعة التجريبية . ودرب تلاميذ المجموعة التجريبية على استخدام بطاقات الحساب . فاستغرق ذلك من مدرسة الحساب حصتين للشرح . وبعد ذلك كل حصة ٧ دقائق منها ٤ دقائق للإجابة عن بطاقة اليوم . أما تلاميذ المجموعة الضابطة ، وهى فصلان ، فكانت حصص الحساب لديهم تسير دون هذا التدريب . وتوحيداً للظروف ، كانت مدرسة الحساب واحدة بين الفصل التجريبى وأحد الفصلين الضابطين .

وبعد شهرين من الدراسة طبق اختبار ستانفورد الجديد فى الحساب مرة ثانية على المجموعتين فتبين أن متوسط المجموعة التى نالت تدريباً على البطاقات أعلى من متوسط فصلى المجموعة الضابطة . والفرق بين المتوسطين جوهرى ، ولكنه غير كبير . وبافتراض أن عدم كبر الفرق الجوهرى يرجع إلى قصر مدة التدريب ، أعيدت التجربة على ثلاثة فصول فى السنة السادسة الإبتدائية أيضاً ، ذلك عند بدء العام الدراسى التالى . وبنفس الطريقة سارت خطوات التجربة . وبعد تدريب ٤ شهور أعيد قياس مستوى الفصول الثلاثة فى الحساب بإعادة تطبيق اختبار ستانفورد ، فتبين أن المجموعة التى دربت على البطاقات فاقت المجموعتين الأخريتين . وكان الفرق جوهرياً وأكبر من الفرق السابق ، مما يدلنا على قيمة هذه البطاقات .

إلا أننا وجدنا تفاوتاً كبيراً بين مستويات تلاميذ الفصل التجريبى ، إذ أن مستوى البعض فى بعض العمليات دل على أنها أعلى من مستوى قدراتهم وتحصيلهم : وأنهم لم يتقنوا بعض العمليات برغم استمرار تدريبهم شهوراً . ومن هنا اتضح لنا أن هناك من يعانون مشاكل القيام بنوع معين

من العمليات وهذه هى الفائدة التشخيصية للبطاقات : وقد أفادت مدرسة الفصل من هذا الكشف بالتركيز على هذه العمليات . ولكن عدم فهم التعليمات أو عدم الارتفاع إلى المستوى المتوسط دل ، كما سبق القول : على احتمال عدم مناسبة مستوى بعض العمليات لهذا المستوى الدراسى .

ولهذا يجرى الآن بحث لتقييمها فى السنة الأولى الإعدادية بإحدى المدارس بنفس الطريقة . وتدل النتائج الأولية على صلاحيتها فى هذا المستوى سواء من حيث القيمة التشخيصية أم التدريبية .

نحو المجلد الثانى

سنتناول فى المجلد الثانى الذكاء والقدرات والاستعدادات والسمات والميول والاتجاهات والقيم و . . . الخ ، ولكل منها أثره ، كما سبق القول ، على التحصيل وعلى التلميذ بما هو فرد . وعلى هذا فكل أدوات تقدير ومساعدة التحصيل والعامل بالتربية والتلميذ ، لم تكتمل بعد : وسنتعرف ، من ثم ، على تكوين الفرد وجوانبه وكيف تقاس . ويمكن للمهتم بالتحصيل وغيره من شئون التربية ، بهذه الطريقة ، أن يفيدها من المجلد الثانى لاستكمال الأدوات والتطبيقات حتى نصل إلى فهم شامل متكامل هادف أفضل .

قائمة مراجع القسم الثالث

1— Adkins, D.C., et al Construction and Analysis of achievement tests. Washington, D.C. : Government Printing office, 1947.

2— Allen, Wendell C., Cumulative Pupil records, New York, Bureau of Publications, Teachers College, Columbia University, 1943.

3— American Psychological Association, Technical recommendations for Psychological tests and diagnostic techniques. Washington : Amer. Psychol-Assoc. 1954.,

4— Anastasi, Anne. Psychological Testing. (2. nd. edition), N.Y. : Macmillan Co., 1957.

5— Anonymous. Tests and Testing programs. Princeton, N.Y. : Educational Testing Service, 1948.

6— Anonymous. Results of Orientation Test in Schools of business of fifty Colleges American of Acountants, College Testing Program, Bulletin No. 4. New York Committee on Selection of Personnel, 1947.

7— Arsenian, S. Informing College Freshmen of their Test Scores. Sch. and Sco., 1942.

8— Barnes, M.W. Gains in The A.C.E. Psychological Examination during The Freshman-Sophomore years. Sch. and Soc., 1943.

9— Baron, D. and Bernard H.W. H.W. Evaluation Techniques for Classroom Teachers, Mc-Graw Hill Book Co., 1958.

10— Barr, A.S., et al. The Measurement and Prediction of Teaching efficiency : A Summery, Inverstigation, J. exp. Educ., 1948.

11— Barrett, D. W. Prediction of achievement. in Typrewiting

and Steneography in a liberal aris College. J. appl. Psychol., 1946.

12— Bean, Kenneth L., Construction of educational and Personnel Tests, New York, McGraw-Hill, 1953.

13— Beatty, J.D, and Cleeton, G.U. Predicting achievement in College and after Graduation. Pers. J., 1928.

14— Bennett, G.K., and Cruikshank, Ruth, M.A. Summary of Clerical Tests N.Y. : Psychol. Corp., 1949.

15— Betts, E.A. Foundations of Reading Instruction. N.Y. : American Book Co., 1946.

16— Betts, E.A. Betts Ready to Read Tests : Manual Meadville, Pa. : Keystone View Co., 1934.

17— Berdie, R.F. Prediction of College achievement and Satisfaction. J. appl. Psychol., 1944.

18— Blackstone, E.G., and McLoughlin, M.W. Blackstone Stenographic Proficiency Tests. Stenography Test : Manual of Directions. Yonkers-on-Hadson, N. Y. : World Book Co., 1932.

19— Blair, G.M., and Kammam, F.F. Do intelligence Tstes requiring reading ability give Spuriously low Scores To Poor readers at the College level ? J. educ. Res., 1942.

20— Blommer, Paul, and dingist, E.F." Rate of Comprehension of reading; its measurement and its relation to comprehension.") educ. Psychol., 1944.

21— Brdin, E.S., and Bixler, R.H. Test Selection : a Process of Counseling. Edec. Psychol Measmt., 1946.

22— Brewer, J.M. Education as guidance New York, : Macmillan, 1932.

23— Brown, Clora M., et al Minnesota Food Score Cards. Minneapolis : Univer. Minnesota Press, 1946.

24— Buros, O.C, (Ed) The Nineteen-Forty Mental Measmure-

ments Yearbook. Highland Park, N.J. : Mental Measmt. Yearbook, 1941.

25 — Buros, O.K. (Ed) The Third Mental Measurements Yearbook. New Bruns - wick, N. J. : Rutgers Univ. Press, 1949.

26— Buros, O.K. (Ed.) Mental Measurement Yearbook, New Brunswick, N.J. : Rutgers Univ. Press, 1938, 1939, 1940, 1949.

27— Burt, C. The Reliability of Teachers Assesment of Their Pupils. B.T. Educ. P. Vol. XV, 1945.

28— Buswell, G.T., and John, Lenore. Diagnostic Studies in arithmetic. Suppl. educ. Monoge., 1926.

29— Buswell. G.T., and John. Lonore. Diagnostic Test For Fundamental Processes in Arithmetic : Manual. Bloomington, Ill. : Public School Publ. Co., 1925.

30 — Chamberlin. D., et al., Did they Succeed in College ?, New York Harper, 1942.

31— Chapman, J.C. Trade Tests, N.Y. : Henry Holt, 1921.

32— Charters, W.W., and Douglas Waples, The Commonwealth Teacher - Training Study, University of Chicago Press, Chicago ; 1929.

33— Columbia Research Bureau Algebra Test 2 : Manual. Yonkers-on-Hadson, N.Y. : World Book Co.. 1933.

34— Conrad, H.S. Summary report on research and development of the Navy's appitude testing program. Final Report on Contract OEM sr 705, OSRD, 1945, Publ. Bd. No. B 289. Washington, D.C., Dept. of Commerce, 1946.

35— Cooperative Algebra Test : Elementary ; Intermediate. Princeton, N.J. : Coop. Test Div., Educ. Testing Service, 1951.

36— Cooperative Contemporary Affairs Test For College Students. Princeton. N.J. : Coop, Test Div., Educ. Testing Service, 1953.

37— Cooperative English Test : Manual. Princeton, N.Y. : Copp. Test Div., Educ. Testing Service, 1953.

38— Cooperative General Culture Test. Princeton, N. Y. : Coop Test Div., Educ. Testing Service, 1951.

39— Cooperative General Achievement Tests : Manual. Princeton, N.Y. Coop. Test Dev., Educ. Testing Service, 1954.

40— Cooperative Mathematics Tests for Grades 3, 8, and 9. Princeton, N.Y. : Coop. Test Div., Educ. Testing Service, 1951.

41— Cooperative Plane Geometry Test. Princeton, N.Y. : Coop. Test Div., Educ. Testing Service, 1951.

42— Cottrell, Donald P. (ed.) Teacher Educatiou for & Free People, The American Association of Colleges for Teacher Education, oneonta, N.Y., 1956.

43— Coxe, W.W, and Cornell, E.L. Prognosis of teaching ability of Studentl in New York State normal Schools, Univ. of State of N.G. Bull. 1934.

44— Cronbach, L.J. Eesentials of Psychological Testing. Harper Broths. N.Y., 1949.

45— Crow, L. E. and Crow, A. Educational Psychology, American Book Co. (Revised ed.) 1958.

46— Darley, J.G., Testing and Counseling in the high-School Guidance Program. Chicago : Science Research Associates, 1943.

47— Davis, F.B. Item Analysis in Relation to Educational and Psychological Testing. Psychol. Bull., 1952.

48— Diam and, Leon N. "Testing the Test makers" Sch. Sci. and math. 1932.

49— Diederich, P.B. The Measurement of Skill in Writing, School Review, 1946.

50— Doob, Leonard W., Public Opinion and Propag. N.Y. : Henry Habt Co., 1948.

51— Douglass, H.R., and Margaret Tallmadge, How university Students prepare for now Types of examinations, Sch. Soc, 1934.

52— Dougss, H R. Different Levels; and patterns of ability necessary For success in College Occupations, 1943.

53— Drought, N.E. An analysis of eight Measures of Personality adjustment in relation to relative Scholastic achievement. J. appl. Psychol., 1938.

54— Dubois, P.H. (Ed.) The Classification Program AAF Aviation Psychology Report. No, 2. Washington, D.C. : Government Printing Office, 1947.

55— Dunlap, J.W. The Predictive value of items for achievement in Various School Subjects. J. appl Psychol., 1935.

56— Durost, Walter N., what constitutes a Minimal testing Pragram ?, Educ. Psychol. Meas., Vol. 7, Spring, 1947.

57— Durost. W.N., et al. Evaluation and Adjustment Series Handbook. Yonkers-on-Hudson, N.Y. : World Book Co. 1937.

58— Durrell. D.D. Durrell Analysis of Reading Difficulty : Manual of Directions. Yonkers-on-Hudson, N.Y. : World book Co., 1937.

59— Dyer, H.S. College Board Scores : Their Use and Interpretarion. No. 1. N.Y. : College Entrance Examination Board, 1953.

60— Dyer. H.S. Evidence on the Validity of the U.S.A.F.I. Tests of General Educational Development. Educ. Psychol. Measmt., 1945.

61— Dysinger, W.S. Test reports in educational Counseling. Educ. Psychol. Measmt., 1943.

62— Ebel, Robert L, Writing the test item, E.F. Lindquist, Editor. Educational Measurement, Washington, D.C., American Council on Education, 1951.

63— Eckert, R., and Marshall, T.O. when youth leave School. New York : Mc Graw-Hill, 1938.

64— Eells, W.G. Scholastic ability of Secondary School Pupils Educ. Rec, 1937.

65— Elsbrea, Willard S., Pupil progress in the elementary School, New York, Bureau of Publications, Teachers College, Columbia University, 1943.

66— Findley, W.J. Agroup testing Program for the madern School. Educ. Psychol. Measmt., 1945.

67— Flanagan, J.C.A Bulletin Reporting the Basic Principles and Procedures Used in the Development of Their System of Scaled Scores. N.Y : Coop. Test Service. 1939.

68— Flanagan, J.C.A Preliminary Study of the 1940 edition on the National Teachers Examination. Sch. and Soc., 1941.

69— Frank W. Hart, Teachers and Teaching, The Macmillan Co., N.Y., 1934.

70— Freeman, F.S : Theory and Practice of Psychological Testing, New York : Henry Holt and Co. Inc. 1950.

71— Froelich, Clifford P., and Arthur L. Benson : Guidance Testing. Chicago : Science Research Associates Inc. 1948.

72— Gardner, E.F. The importance of reference groups in Scaling Procedure. Proc. 1952.

73 Gardner, E.F. Value of norms bosed on a new type of Scale unit. Proc., 1948.

74— Garretson, O.K. Relationships between expressed preferences and Curriculom abilities of ninth grade boys. New York : Teachers College Contributions to Education, No. 396, 1930.

75— Gates, A.I. Gates Basic Reading Tests : Manual of Directions. N.Y. : Teachers coll., Columbia Univer., Bur. Publ., 1943.

76— Gaudet, F.J., and Marryott, E.S. Predictive value of the

Ferson-Stoddard Law Aptitupe Examination. An. Law Sch. Rev., 1930,

77— Gerberich, J.R. Validation of a State-wide educational guidance Program for high-School Seniors. Sch. and Soc., 1931.

78— Ghiselli, E.E., and Brown, C.M. Personnel and Industrial Psychology. N.Y. : Mc Graw-Hill, 1948.

79— Good, C.V. and Merkel W.R. Dictionary of Education. (2 nd edition) Mc Graw-Hill, N.Y. 1959.

80— Goodenough, Florence L. Mental Testing, Rinehart, N.Y. 1949.

81— Gotham, A.E. Personalily and teaching efficienny. J. exp. Educ. 1945.

82— Gray, W.S. Standardized Oral Reading Paragraphs : Manual. Bloomington, Ill. : Public School Publ. Co., 1915.

83— Greene, E.B.A. Sompling of Vocabularies of Superior adults. J. higher Educ., 1933.

84— Greene, E.B. Michigan Vocabulary Profile Test : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1949.

85— Greene, E.B. The Wichigan Vocabulory Profile Test, after ten years. Educ. Psychol. Measmt., 1951.

86— Greene, H.A., and Bruce, H.W. Iowa Plane Geometry Aptitude Test : Iowa City : Bur. Educ. Res. Service, Univer. Iowa, 1942.

87— Greene, H.A., Jorgensen, A.N., and Kelley V.H. Iowa Silent Reading Tests, New Edition (Revised). Advanced Test : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1943.

88— Greene, H.A., and Kelley, V.H. Iowa Silent Reading Tests New Edition (Revised). Elementary Test : Manual of

Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1943.

89— Greene, H.A., Wright, F.B., Ruch, G.M., and Studebaker, J.W. Compass Survey Tests in Arithmetic : Manual. Chicago : Scott, Foresmae, 1927.

90— Greenc, H.A. and Jorgenson A.N. and Jerberich, J.R. Measurement and Evaluation in The Secondary School. Longmans Grean and Co. N.Y. 1947.

91— Greene, E.B. Measurement of Human Behavoir. Odessey N.Y. : 1941.

92— Gregory, W.S. From high School to College. Occupations, 1939.

93— Gregory, W.S. Data regarding The Academic Interest Inventory. Educ. Psychol. Measmt., 1946.

94— Guilford, J.P. New Standards For Test evaluation. Educ. Psychol. Measmt, 1946.

95— Gulliksen, H. Intrinsic Validity. Amer. Psychologist 1950.

96— Guliksen, H., Theory of Mental Tests, 1950. & N.Y. : John Wiley & Sons.

97— Guilford, J.P. Psychometric Methods. N.Y. : Mc Graw-Hill, 1936.

98— Guilford. J. P. Factor analysis in a Test-development Program. Psychol. Res., 1948.

99— Harry, D.P., Jr., and Durost, W.N. Essential High School Content Battery : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1951.

100— Hartmann, G.W. Measuring Teaching efficiency among College instructors. Arch. Psychol., No. 154, 1933.

101— Hawkes, Herbert, et al (eds.) The Construction and Use of Achievement examinations. Boston : Houghton. Miffin, 1936.

102— Hessemer, M. The Thurstone PMA Tests in a study of academic success With School of engineering chemistry and physics. Unp-ubl. thesis, penn. State College, 1942.

103— Hildreth, Gertrude H., Metropolitan achievement tests : manual for interpreting, Yonkers. New Yoıc, world Book., 1948.

104 — Hilolreth, Gertrude H., and Griffiths, N. L. Metropolitan Readiness Tests : Directions for Administrating and Scoring. Yorkers-on-Hudson, N. Y.: woıld Book Co., 1949.

105— Hillegas, M. B. A. Scale for the measurement of quality in English Composition by young People. Teachers coll. Rec., 1912.

106 — Jones. G., and Galbraith, A. The interpretation of standardired tests. Sch. and Soc., 1941.

107 — Jones, R. D. Prediction of teaching efficiency From objective measures. J. exper. Educ., 1946.

108 — Jordan, A. M. Educational psychology. New York: Henry Holt, 1928.

109 — Jordan, A. M. : Measurement in Education, N. Y. : Mc. Grow Hill Book Co., 1953.

110 — Kawin, Ethel, Records and reports : obeservations, tests, and measurements, Early Childhood education, Forty-Sixth Yearbook, National Society for the study of Education, 1947.

111 — Kelley, T. L., and Ruch, G. M. Stanford Achievement Test : Manuals for primory, Elementary, Intermediate, and Advanced Batteries. Yonkerson - Hudson, N. Y. : World Book Co., 1953

112 — Kelley, T. L. Ridye-route norms. Harvard educ. Rev., 1940

113 — Kelley, T. L. and Kery, A. C. Tests and Measeremens in the Social Sciences. N. Y. : Scribner, 1934.

114 — Kohn, H. A. Achievement and intelligence examinations

Correlated with each other and with teacher's rankings. J. genet. psychol., 1938.

115— Lee, J. M. Lee Test of Algebra Ability : Manual. Bloomington, Ill. : Public School publ. Co., 1930.

116— Lee, Doris M., and Lee, J. M. Lee Test of Geometric Aptitude : Manual. Los Angeles : California Test Bur., 1931.

117— Lewis, E. E. Lewis Scale for Measuring Special Types of English Composition. Yonkers-on-Hudson, N. Y. : Wsrld Book Co., 1921.

118— Lindquist, E. F., Iowa tests of General Educational Development : Manual, Chicago, Ill., Science Research Associates, 1951.

119— Lindquist, E. F. (Ed.) Educational Measurement. Washington : Amer. Council Educ., 1951

120— Lindquisb, E. L. (Ed.) Meesurement Handbook. Washington, D. C. American Council on Education, Boston : Houghton Mifflin Co., 1951

121— Lindquist, E. F., Preliminary Considerations in objective test Construction, Chapter 5 in E. F. Lindquist, Editor, Educational measurement, Washington, D. C., - American Council on Educatisn, 1951.

122— Little, J. K. Result of use of machines For testing and For drill, upon Learning in educotional psychology. J. exp. Educ., 1934.

123— Lord, A B A. Survey of Four hundred Forty-nine special Class Pupils. J. educ. Res., 1933.

124— Lorge, I. The Fundamental Nature of Measurement, In. E. F. Lindquist (ed.) Educational Measurement. Washington : Amer. Council Educ., 1951.

126— Mathews, L. H. An Item Analysis of Maasures of teaching ability. J. Educ. Res., 1940.

125— Mc. Call, W. A., How to Measure in Education, 1922

126— Meyer, G. An experimental study of old and new types of examination, J. educ. Psychol., 1934.

127— Meyer (george. "The effect on Recall and Recognition) of the examination set in classroom Situations" G. Educ. Psychol., 1937.

128— Micheels, William J., and M. Ray Karnes. Measuring educational achievement, New York, Mc. Graw-Hill, 1950

129— Monroe, Walter S., Editor Encyclopedia of educational research, reved., New York, Macmillan, 1950

130— Monroe, Marion. Children who Can not Read. Chicago : Univer. Chicago Press, 1932.

131— Monroe, Marion. Reading Aptitude Tests : Manual. Boston : Houghton Miffin, 1935

132— Mook E, H. Psychology For business and industry. New York : Mc Graw-Hill, 1942.

133— Mozzei, R. Desirable Traits of Successful Teacher, J. Teach. educ., 1951.

134— Mursell, J. L. Psychological Testing. (2 nd edition) Longmans Green and Co. 1950.

135— National Society for The Study of Education, The Measurement of understanding, The Forty-fifth Year book, Part I, Chicago, Illinois, University of Chicago Press, 1946.

136— Nunally Jum C. Tests and Mearurements Assessment and Prediction. Mc-Graw-Hill Book Co. N.Y. 1959.

137— Odell, C. W., How to improve classroom Testing, Dubuque, Lowa. William C. Brawn, 1953.

138— Orleans, J. B., and Orleens, J.S. Orleans Geometry Prognosis Test, Revised Edition : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1951.

139— Orleans, J.B., and Orleans, J.S. Orleans Algebra Prog-

nosis Test, Revised Edition : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co. 1951.

140— Patterson, C.H. on The Problem of the Criterion in Prediction Studies. J. Consult. Psychol., 1946.

141— Paterson, D.G., Schnelidser, G. and Williamson, E.G. Student guidance Techniques. New York : Mc Graw-Hill, 1938.

142 – Paul A. Witty, The Teacher Who Has Helped the Most, National Educ. Assoc. J. 1947.

143 – Payne, A.F. Sentence Completions New York Guidance Clinic, 1928.

144— Pond. M. Occupations, intelligence, age and Schooling. Pers. J.. 1933.

145— Pressey, Luella G. An investigation of The Tecnical Vocabularies of the School Subjects, Educ. Res. Bull., Ohio State Univer., 1924.

146— Raths, lovis, "Basis four Comprehensive evaluation" Edic. Res. Bull, 1936.

147— Remmers, H.H. and Gage, N.L. and Rummel, J.F. Practical Introduction to Measurement and Evaluation. Harper Brothers 1960.

148— Rice, J.M. Futility of Spelling Grind, Forum, 1897.

149— Rice, J.M. The Public School System of the United States, N.Y. : The Century Co., 1893.

150— Richey, R.W., and William H. Fox. "A Study of Some Opinions of High School Students with Regard to Teachers and Teaching" Bulletin of the School of Educ Indiana University Bloomington Ind., 1951.

151— Richey, R.W. Planning for Teaching (2 nd.) N.Y. : Mc Graw Hill Co, 1958.

152— Rosenfeld, M.A., and Nemzek, C.L. Long-range Prediction of College marks. Sch. and Soc., 1938.

153— Ross, C.C. Should Low-ratinh Dollege Freshmen be told their Scores in intelligence Tests? Sch ond Soc., 1938.

154— Ross, C. C. Measurement in Today's Schools. N.Y. : Prentice-Hall, 1947.

155 — Ruch, G,M,, et al, Compass Diagnostic Tests in Arithmetic : Manual Chicago : Scott, Faresmon, 1925,

156 — Rukh, N,M,, and Stoddard, G,D. Comparative Reliabilities of Five Types of Objective Examinations, J. Educ, Psychol,, 1915,

157— Ruch, G,M,, and Stoddard, G,D, Tests and Measurement in High School Instruction, N,Y, : World Book Co,, 1927.

158— Russell, B, History of Western Philosophy, London, George Gllen and Unwin, 1945,

159— Schneidled, G,G., and Berdie, R,F, Educational ability Patterns, J, educ, Psychol,, 1942,

160 — Schneid ler, G,G, Grade and age norms For the Minnesoia Vocational Test For Clerical Workers. Éduc, Psychol, Measmt., 1941.

161— Seagoe, M,V, Prognostic tests and Teaching Success, J, educ, Res., 1945,

162 — Seagoe, M,V, Permanence of interest in Teaching, J, educ. Res,, 1945.

163 — Seashore, H,, and Bennett, G, K, Seashoce-Bennett Stenographic Proficiency Test : Wanual, N,Y, : Psychol, Corp,, 1946,

164 — Seder, M, The reliability and Validity of the A,C,E, Psychological Examination, 1938,

165 — Segel, David, State Testing and evaluation Programs, Federal Security Agency, Office of Education, Circular No, 320, Washington, D,C, U,S, Office of Education, 1951,

166– Shaffer, L,F, Informotion Which Should be Provided by Test Publishers and testing agencies on the Validity and use of their tests : Personality tests, Proc 1949 Invrit Conf testing Services,

167–Skinner, C. E. Educational Psychology. Prentice Hall 1950

168– Smith, E. R., Tyler, R. W., et al Appraising and Recording Student progress. N.Y. : Harper, 1942

169– Sorenson, H. Psychalagy in Education N. Y., 1954

170– Spitrer, H. F., et al. Iowa Every-pupil Tests of Basic Skills : Manual Boston : Houghton Miffin, 1947

171– Staff, Personnel Research Section, A.G. O. Validation of the General Clerical Abilities Test for Selection and Placement of War Dept., Civilian Personnel. Amer. Psy, 1947

172– Stalnaker, John M., The essay type of examination, Chapter 13 in E. F. Lindquint, Editor, Educational measwement, Washington, D. C. American Council on Dducation, 1951

173– Stead, W. H., Shartle, C. L., et al · Occupational Covnselinh Techniques. N. Y. : American Book Co., 1940

174– Ttephens, J.M. Educotional Psychalagy N. Y, 1951

175– Stevens, S. S. On The Theory of Scales of Measurement Science, 1946

176– Stopher, E. C. The freshmen testing program. J. higher Educ., 1941

177– Strang, R. Behavior and backgronnd of students in college and Secondary School New York : Harper & Brothers, 1937

178– Strang, R. Personal development and guidance in college and secodary School. New York : Harper & Brothers,1934

179– Strang, Ruth, Reporting to parents, New York Bureau of publications, Teachers college, columbia University, 1947

180 — Stuit, D. B. Astudy of the vocational interests of a group of teachers, college freshmen. J. oppl. Psychol., 1938

181— Stuit. D. B. (Ed.) Personnel Research and Test Develo-pement in the Bureau of Naval personnel. princeton, N. J. : Princeton Univer. Press, 1947

182— Super, D. E. Appraising Vocational Fitness by means of Psychological Tests. N. Y. : Harper Brothers, 1949

183— Tiegs E. W., and Clork, W. W. California Achievement Tests : Manuals of Directions For Primary, Elementary, Intermediate, and Advanced Batteries. Los Angeles : California Test Bvr., 1951

184— Terry, P. W., How students study for three typs of objective tests, J. educ. Res., 1934

185 — The Evalualion of Studeut Teachlng, Twenty-eighth Annual Yearbook of the association for Sindent Teaching State Teachers College, Lock Haven, 1949

186— The Purposes of Education in American Democracy, National Education Associations, Educational Policies Commission. Washington, 1938

187— Thomlinson, Loren R. Recent Studies in the evaluation of teaching, educ. Res - Bull., 1955

188— Thorndike R. L. and Hagen, A. Measurement and Ev ation in Psychology and Education. John Willy Inc. N. Y. 1957

189— Thorndike, R. L., and Lindquist, E. F. Educational Measurement, 1951

190— Thorndike R. L. (Ed.) Research Problems and Techni-ques. AAF Aviation Psychology Report, No. 3 Washington D. C. : Government Printing Office, 1947

191— Thurstone. L. L, A, Method of Scaling Psychological and Educational Tests, J, Educ, Psychology, 1925

192— Thurstone, L, L, Scales Construction With Weighted Observation, Psy., 1928,

193— Thurstone, L.L. The Unit of Measurement in Education Scales, J.E. Psy., 1927.

194 - Travers, Robert M.W., How to make achievement tesl New York, odyssey Press, 1950.

195— Travers, R.M.W., Craig, H. and Fenchel, G. Objectiv of Teacher Education, J. Educ Res., 1951.

196— Travers, R.M.W. Educational Measurement, N.Y. ma millan Co., 1955.

197— Traxler, Arthur E., Administering and Scoring Tl objective Test, Chapter 10 in E.F. Lindquist, Editc Educational measurement, Washington, D.C., America Council on Education. 1951.

198— Traxler, A. E. Cumulative Test records : Their natu and uses. Educ. Psychol. Measmt., 1941.

199— Traxler, A. E. Evaluation of aptitude and achievement a guidance Program. Educ. Psychol. Measmt., 1946,

200— Traxler, A,E., et al, Introduction to Testing and the U of Test Results in Public Schools, N,N, : Harper, 1953

201— Traxler, Arthur E,, Techniques of guidance, New Yor Harper, 1945,

202— Traxler, Arthur E, The nature and use of reading Test Chicago : Science Research assoc 1941,

203— Traxler, Arthur E, The use of Test results in diagnos and instruction in the tool Subjects, rev, ed,, Education Records Bulletin No, 18, New York, Education Records Bureau, 1949,

204— Triggs, Frances O,, et al, Diagnostic Reading Tes A History of Their Con-Struction and Validation, N,Y, Common Diagnostic Reading Tests, Inc., 1952,

205 — Triggs, Frances O,, et al, Diagnostic Reading Tests Their Interpretation and Use in the Teacing of Rea ing, N,Y, : Comm on Diagnostic Readieg Tests, Inc 1952,

206— Trimble, H,C·, and Cronbach, Lee g. A, practical procedure for the Rigorous interpretation, of test - Retest scores in terms of pupil growth,, J Edue. Res,, 1943,

207 — Troyer, M, E,, and Angell, G, W, Manual for the SRA Self-Scorer, Chicago : Science Research Associates, 1949

208— Trump, J,B,, and Haggerty, Helen R, Basic principles in achievement test item construction P R S Report 979, Personnel Res, Sect, AGO, 1952,

209— Tyler, Ralph W,, The function, of measurement in improving instruction, Chapter 2 in E, F, Liudquist, Editor, Educational measurement, Washingion, D, C,, American Council on Education, 1951,

210— Tyler, R,W, Permanency of Learning, J, higher Educ,, 1933

211— Tyler, R,W,, Constructing acheivement Tests, columbus, Ohio, Bur Educ Res,, Ohio Seote Univ., 1934,

212— United States Armed Forces Institute, Tests of General Educational Developement : Examiner's Manual, Washington : Amer, Council Educ,, 1944,

213— United States office of Education, National Committee on Cumulative Records, Hand book of Cumulative Records, Federal Security Agency, office of Education, Bulletin 1944. No 5, Washington, United States Government Printing office, 1945

214— Van Wagenen, M, J, van Wagenen English Composition Scales, Yonkers-on-Hudson, N.Y. : World Book Co,, 1923,

215— Voughn, K.W, Planning the objective test, chapter 6 in E,F, Lindquist, Editor, Educational meosurement, Washington, D,C, American Council on Education, 1951,

216— Vernon, P.E, Educational abilities of training College Students, Brit J, educ, Psychol,, 1939

217— Warren H,C. Dictionary of Psychology Hanghton Mifflin Co. 1934

218— Watson. G., and Glaser, E. M. Watson-Gloser Critical Thinking Appreisal : Manual. yonkerson Hudson, N. Y. : World Book Co., 1952.

219— Webster, New International Dictionary, 1945.

220— Weitzman, E, and Monamara, W. J. Constructing Classroom Examinations Chicago : Science Research Associates, 1949.

221— Weitrzman, Ellis. and Walter J. McNamare, Constructing classroom examinations : a guide for teachers, Chicago, Science Research Associates, 1950.

222— Wert, J.F. Twin examination assumptions. J. higher Educ., 1937.

223— Williamson, E.G. How to counsel Students. New Yoik : Mc Grow-Hill, 1939.

224— Wilson, M. H. The Self-appraisal Program in the Philadelphia Junioc High Schools, Educ. Psychol. Measmt., 1946.

225— Wilson, J.W., and Carpenter, E.K. The need for restandardization of altered tests, Amer. Psychol., 1948.

226— Witty, Paul A, Mental Hygiene in modern Education, Fifty-fourth Yearbook of the National Society for the Study of Education, University of Chicaho Press Chicago, 1955.

227— Wolf, R.R., Jr. Differential forecasts of achievement and Their Use in educational counseling, Psychol. Monog., No. 1, 1941.

228— Wrightstone J.W. Test of Critical Thinking in the Social Studies : Manual, N. Y. : Teachers Coll., Columbia Unive., Buc. Publ., 1939.

229— Wrinkle, William L., Improving making and reporting Practices in elementary and Secondary Schools, New York, Rinehart, 1947.

230— Yeager, T.C. An analysis of certain of Selected high School Seniors interested in Teaching, New York : Teachers College Contribution, No. 660, 1935.

٢٣١ – أحمد زكى صالح : بطاقة تقويم المدرس : مكتبة النهضةالمصريه سنة ١٩٥٩ .

٢٣٢ – أحمد زكى صالح : علم النفس التربوى « الطبعة الخامسة » مكتبة النهضة المصريه سنة ١٩٥٧ .

٢٣٣ – ارثرجيتس وآرثرت جيرسلد ، ت . ر . ماكونل ، وروبرت س . تشالمان : علم النفس التربوى « ٣ أجزاء » ترجمة إبراهيم حافظ ، والسيد محمد عثمان ، ومحمد عبد الحميد أبو العزم بأشراف الدكتور عبد العزيز القوصى .

٢٣٤ – إسماعيل محمود القبانى ، دراسات فى مسائل التعليم : مكتبة النهضة سنة ١٩٥١ .

٢٣٥ – إسماعيل محمود القبانى : دراسات فى تنظيم التعليم : مكتبة النهضة المصرية سنة ١٩٥٨ .

٢٣٦ – إسماعيل محمود القبانى : التربيـــة عن طريق النشاط : مكتبة النهضة المصرية سنة ١٩٥٨ .

٢٣٧ – إسماعيل القبانى ومحمد عبد السلام أحمد : أختبار استانفورد الجديد فى الحساب (الطبعة الخامسه) لجنـــة التأليف والترجمه والنشر سنة ١٩٦٠ .

٢٣٨ – إسماعيل القبانى ومحمد عبد السلام أحمد : البطاقات التدريبية والتشخيصية فى الحساب : لجنة التأليف والترجمه والنشر .

٢٣٩ – الأهداف والمستويات فى التربيه والتعليم : وزارة التربيه والتعليم سنة ١٩٥٥ .

٢٤٠ – جليفورد وآخرون : ميادين علم النفس : أشرف على جمعه جليفورد وعلى ترجمته الدكتور يوسف مراد – دار المعارف سنة ١٩٥٦ .

٢٤١ – رابطة التربية الحديثة : مشكلة الإمتحانات فى مصر : القاهرة : مطبعة دار التأليف والترجمه والنشر سنة ١٩٣٨ .

٢٤٢ – رمزية الغريب : إختبار الاستعداد للمهن التعليمية : كلية البنات سنة ١٩٥٧ .

٢٤٣ – سعد جلال : التوجيه النفسى والتربوى والمهنى : مكتبة النهضه المصرية سنة ١٩٥٧ .

٢٤٤ – عطيه محمود هنا : التوجيه التربوى والمهنى مكتبة النهضه المصريه سنة ١٩٥٩ .

٢٤٥ – محمد عبد السلام أحمد : تقرير بحث تقييم التلميذ بالمرحلة الابتدائية فى المناطق المختلفة لجنة التأليف والترجمة والنشر سنة ١٩٥٦.

٢٤٦ – محمد عبد السلام أحمد : تقرير بحث تقييم امتحانات القبول بالمرحله الاعداديه سنة ١٩٥٥ .

٢٤٧ – محمد عبد السلام أحمد : تقرير بحث تتبع خريجى التعليم الصناعى لجنة التأليف والترجمة والنشر سنة ١٩٦٠ .

٢٤٨ – محمد عبد السلام أحمد : تقرير بحث أثر البطاقات التدريبية والتشخيصية فى الحساب : لجنة التأليف والترجمة والنشر سنة ١٩٦٠ .

٢٤٩ – محمود رشدى خاطر : سلسلة اختبارات سرس الليان فى القراءه للكبار سنة ١٩٦٠ .

٢٥٠ – محمود رشدى خاطر : سلسلة اختبارات سرس الليان للقراءة الصامته سنة ١٩٦٠ .

٢٥١ – مصطفى فهمى : سيكولوچية التعلم (الطبعة الثانية) لجنة النشر للجامعين سنة ١٩٥٣ .

٢٥٢ – وزارة التربية والتعليم : دليل الإختبارات للقبول بالمرحله الاعداديه بإشراف محمد عبد السلام أحمد سنة ١٩٥٥ .

٢٥٣ – وزارة التربية والتعليم : تقرير لجنة بحث التعليم الابتدائى . سنة ١٩٥٧ .

٢٥٤ – وزارة التربيه والتعليم : التقرير السنوى لإدارة البحوث الفنية والمشروعات ١٩٥٩ – ١٩٦٠ .

٢٥٥ – وزارة التربيه والتعليم : دليل تقويم التلميذ فى المرحلة الابتدائية سنة ١٩٦٠ .